PHYS 250

Lecture 5.1-B

Applications 1-B:

Bonds, Bands & Semiconductors

Today

Review
Semiconductor Devices

Review 1

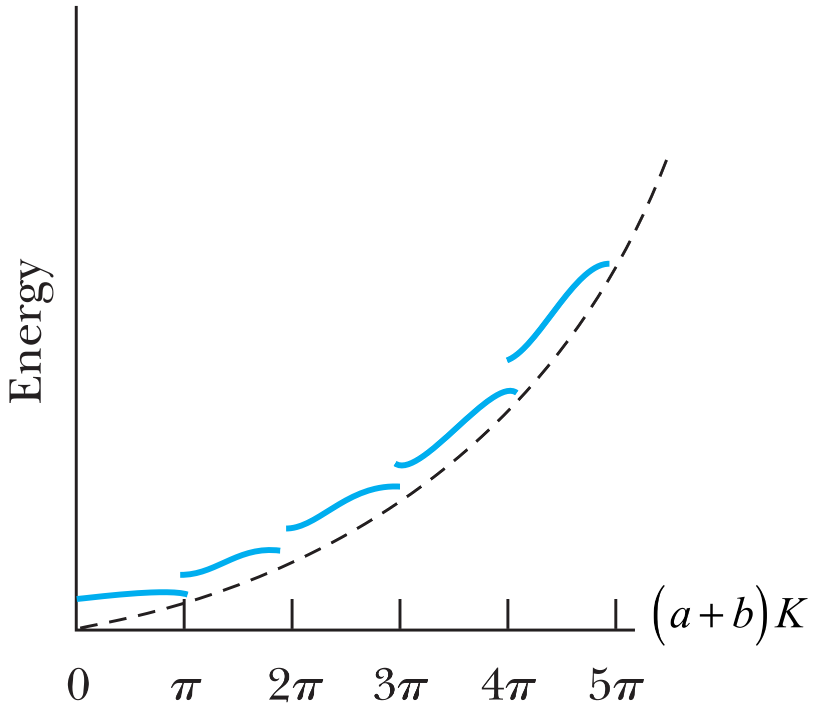
A single potential well has multiple energy levels.

If two potential wells are close together, the energy levels split in two, one going up, one going down.

A particle put into one potential well can tunnel into the other well, and then tunnel back.

If there are N potential wells, each energy level splits into N energy levels. Particles can tunnel between all the wells.

If N is very large, the bound-state energy levels become nearly continuous, but there are still band-gaps between them.



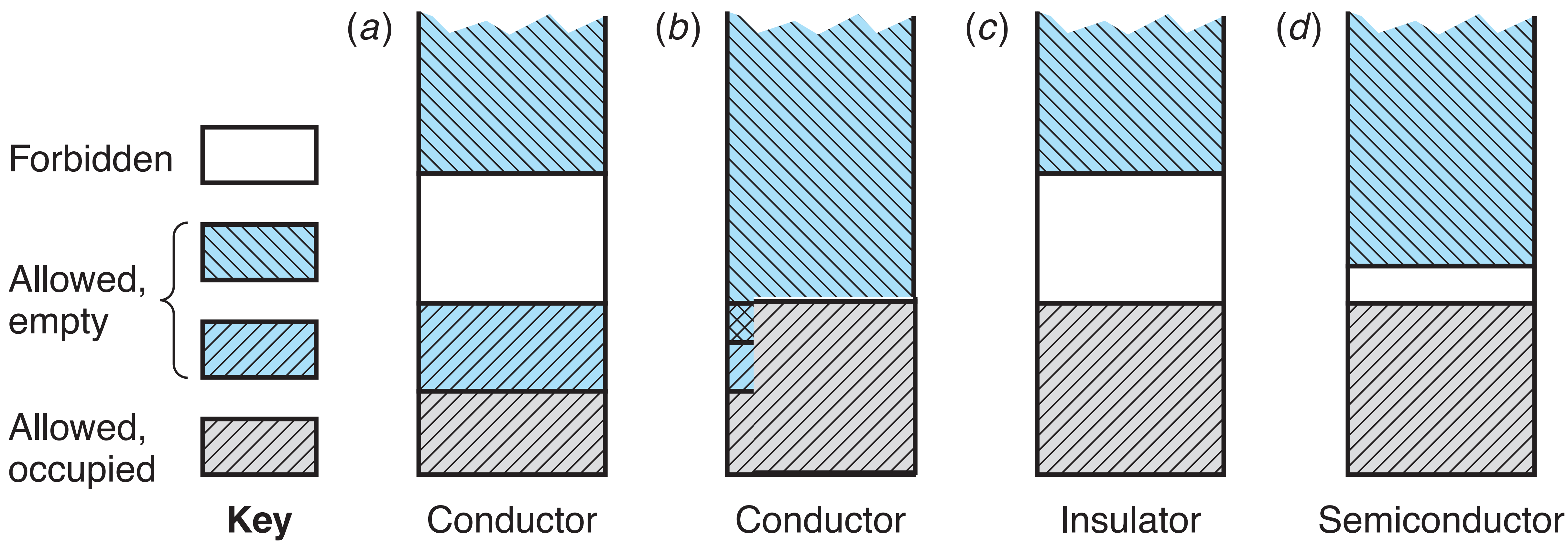
Review 2

The Pauli Principle of 2 electrons per "state" still applies.

A conductor has either a half-filled band, or overlapping bands.

If all the states in a band are occupied, electrons can't move, so no conductivity. If a band is empty, no conductivity. This is an insulator.

A semiconductor is an insulator with a very narrow band-gap.



Conductivity in Semiconductors

The highest occupied band is called the valence band,
and the lowest un-occupied band is called the conduction band.

In a (pure) semiconductor (and an insulator), the valence band is full,
and the conduction band is empty, except a few thermally excited electrons.

If an electron somehow gets into the conduction band, it can tunnel
from atom to atom, and there is conductivity.

If an electron is somehow missing from the valence band, that is called a hole.

An electron from a neighboring atom can tunnel across and fill the hole.
But that leaves a hole at the neighbor atom.

An electron can tunnel across from the next neighbor and fill that hole.

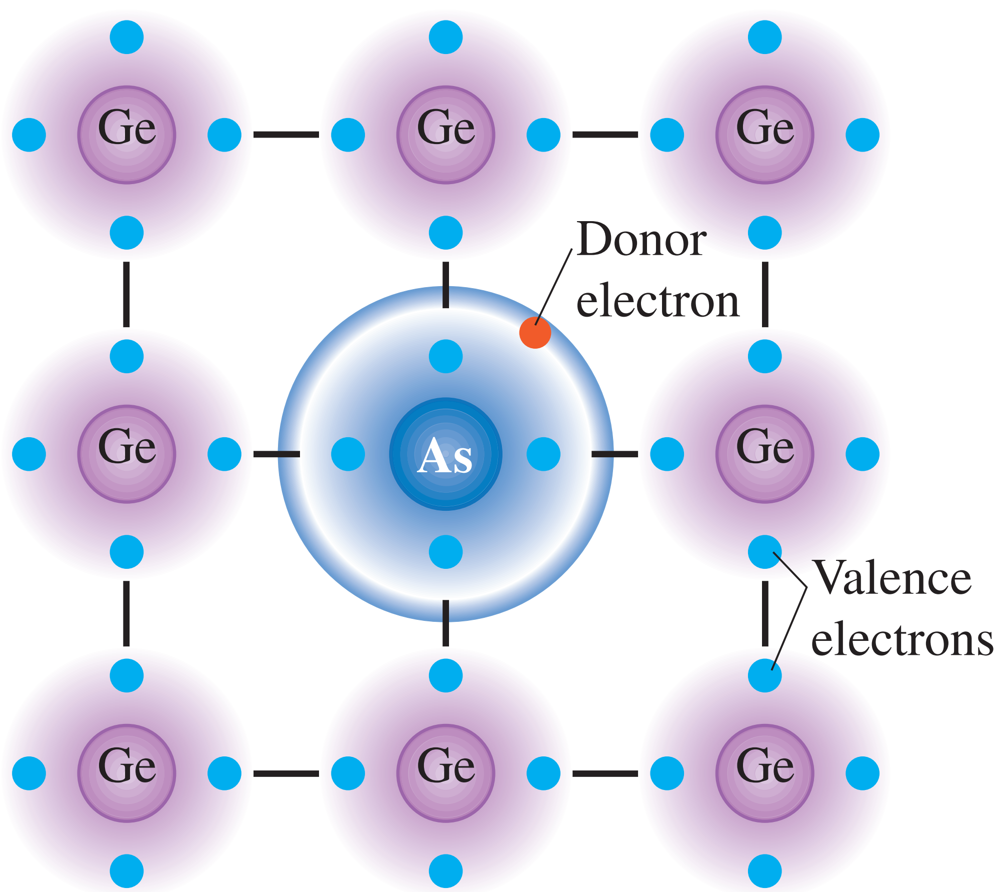
The motion of the hole also contributes to conductivity.

Donor Doping

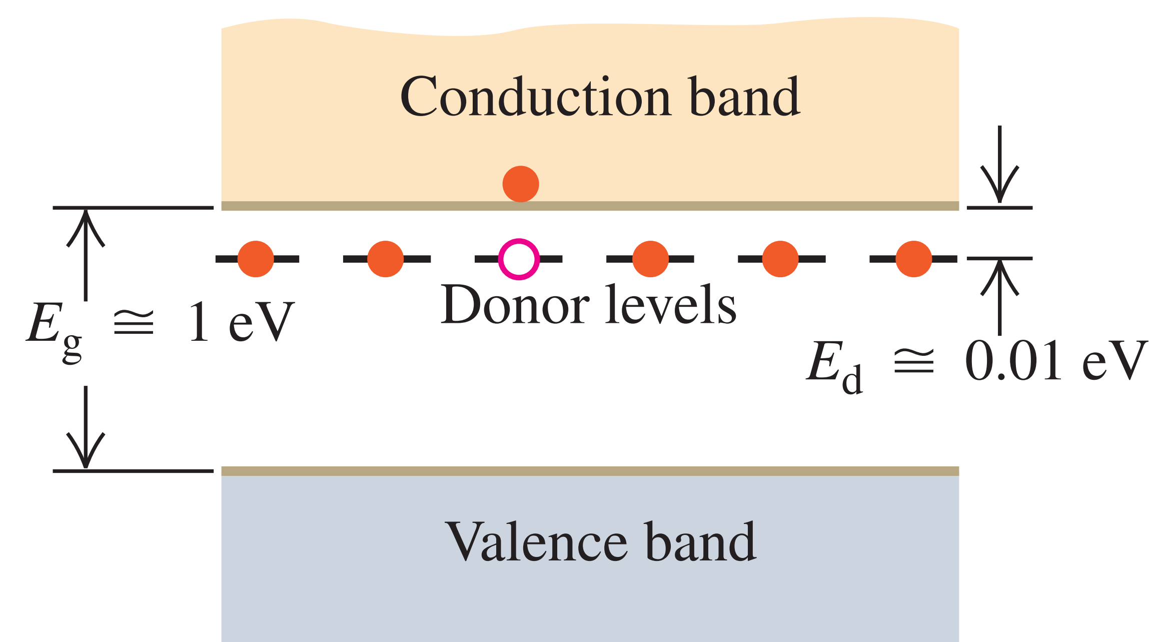
Arsenic has almost the same size as Germanium, and has 5 outer electrons, vs 4 for germanium or silicon.

Replacing N atoms of Germanium by Arsenic puts N extra electrons into the lattice (balanced by N extra charges from immobile Arsenic atoms).

The "donor levels" are very close to the conduction band, so the donor electron is easily thermally excited into the conduction band.



(b) Energy-band diagram for an *n*-type semiconductor at a low temperature. One donor electron has been excited from the donor levels into the conduction band.

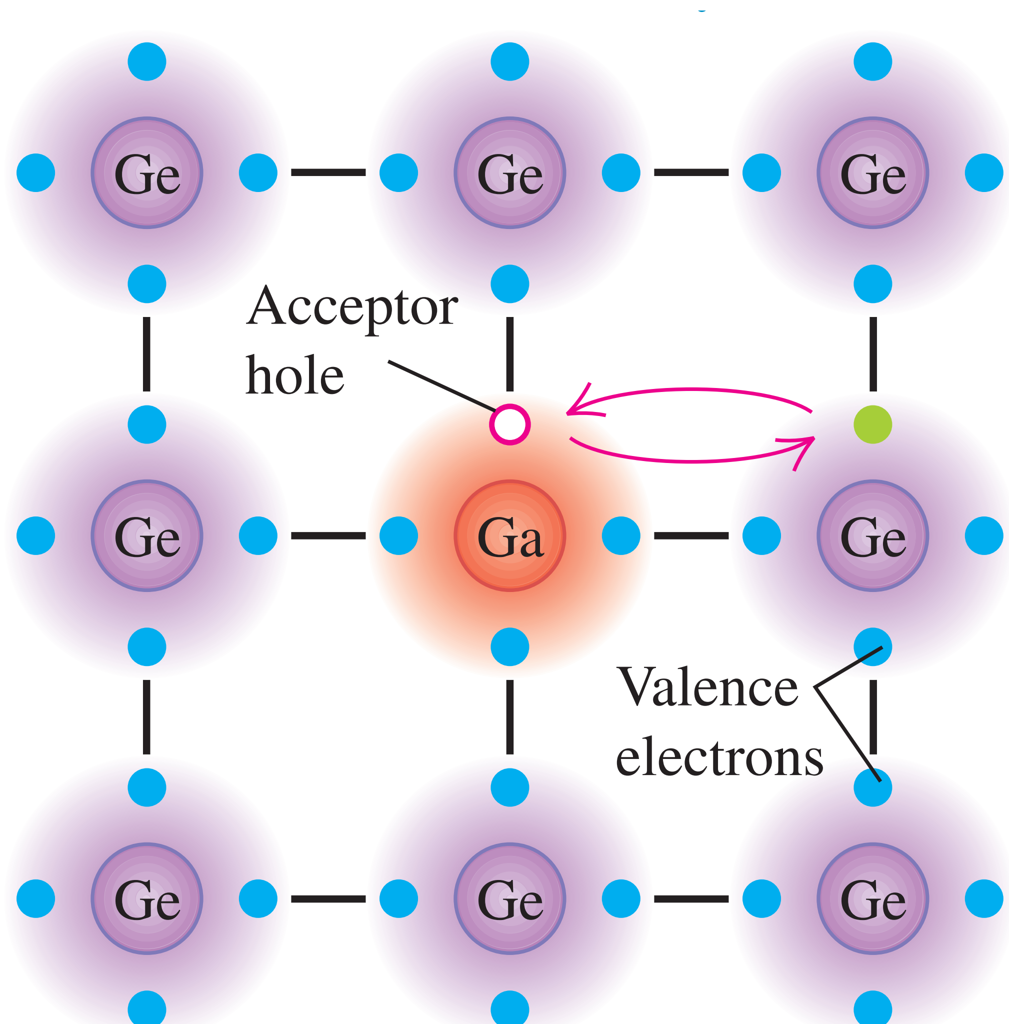


Acceptor Doping

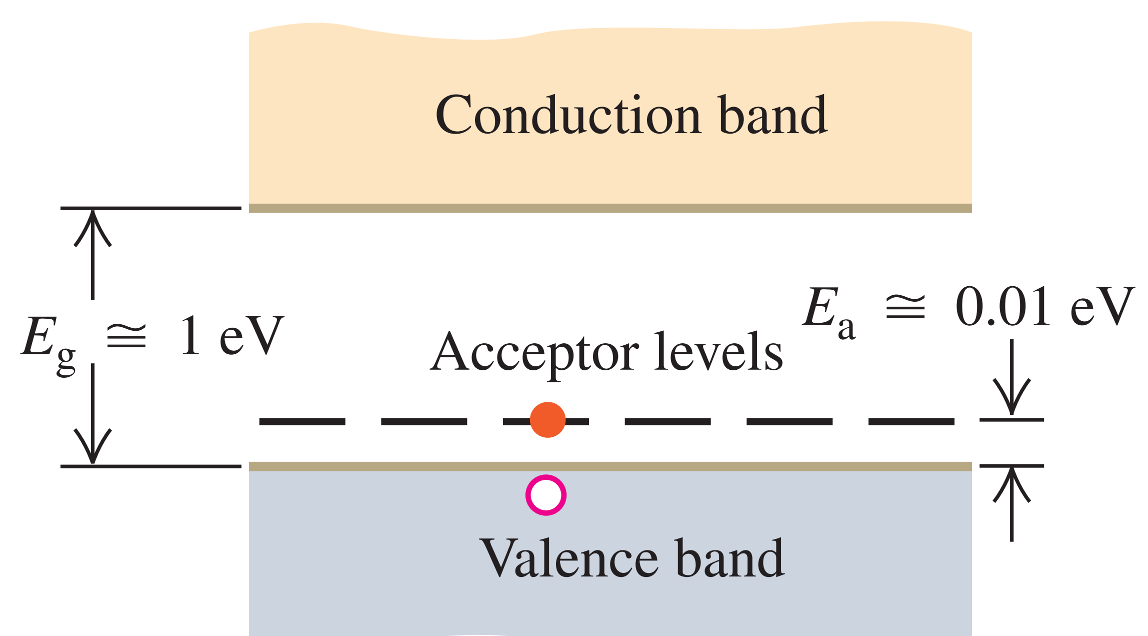
Gallium is almost the same size as Germanium, but has 3 outer electrons.

Replacing N atoms of Germanium with Gallium makes the lattice have N fewer electrons (balanced by immobile Gallium atoms with fewer protons).

The "acceptor levels" are very close to the valence band, so a valence-band electron can jump to a Gallium atom, leaving a mobile hole in the valence band.



(b) Energy-band diagram for a *p*-type semiconductor at a low temperature. One acceptor level has accepted an electron from the valence band, leaving a hole behind.



N-Type and P-Type

Doping with electron donors (5 valence electrons) makes what is called N-type material, which has mobile electrons and fixed positive charges.

Doping with electron acceptors (3 valence electrons) makes P-type material, which has mobile holes and fixed negative charges.

Both N-type and P-type materials have zero net electric charge.

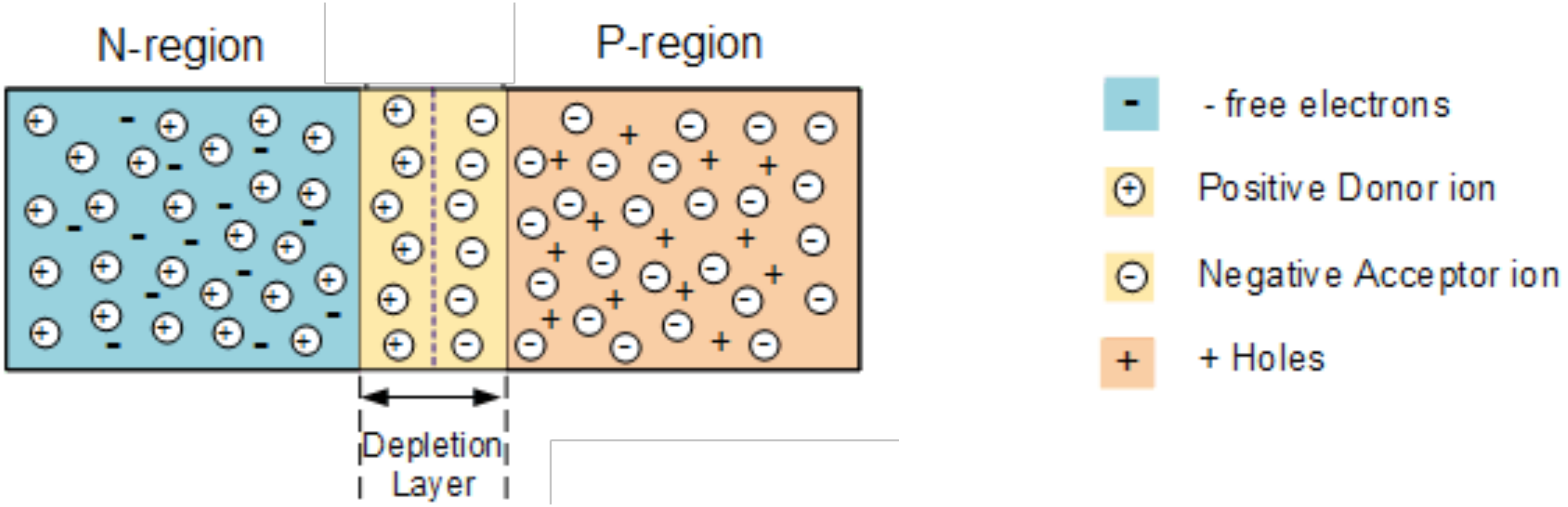
It's best to dope with atoms that are close to the same size as the lattice, which happens when the doping atoms are in the same periodic table row.

Phosphorous doping makes N-type Silicon. Boron doping makes P-type silicon.

Arsenic doping makes N-type Germanium. Gallium doping makes P-type.

PN Junction

Start with a pure silicon wafer. Diffuse some P-type impurity into the whole thickness. Then diffuse enough N-type impurity to reverse the polarity, but just of the surface.

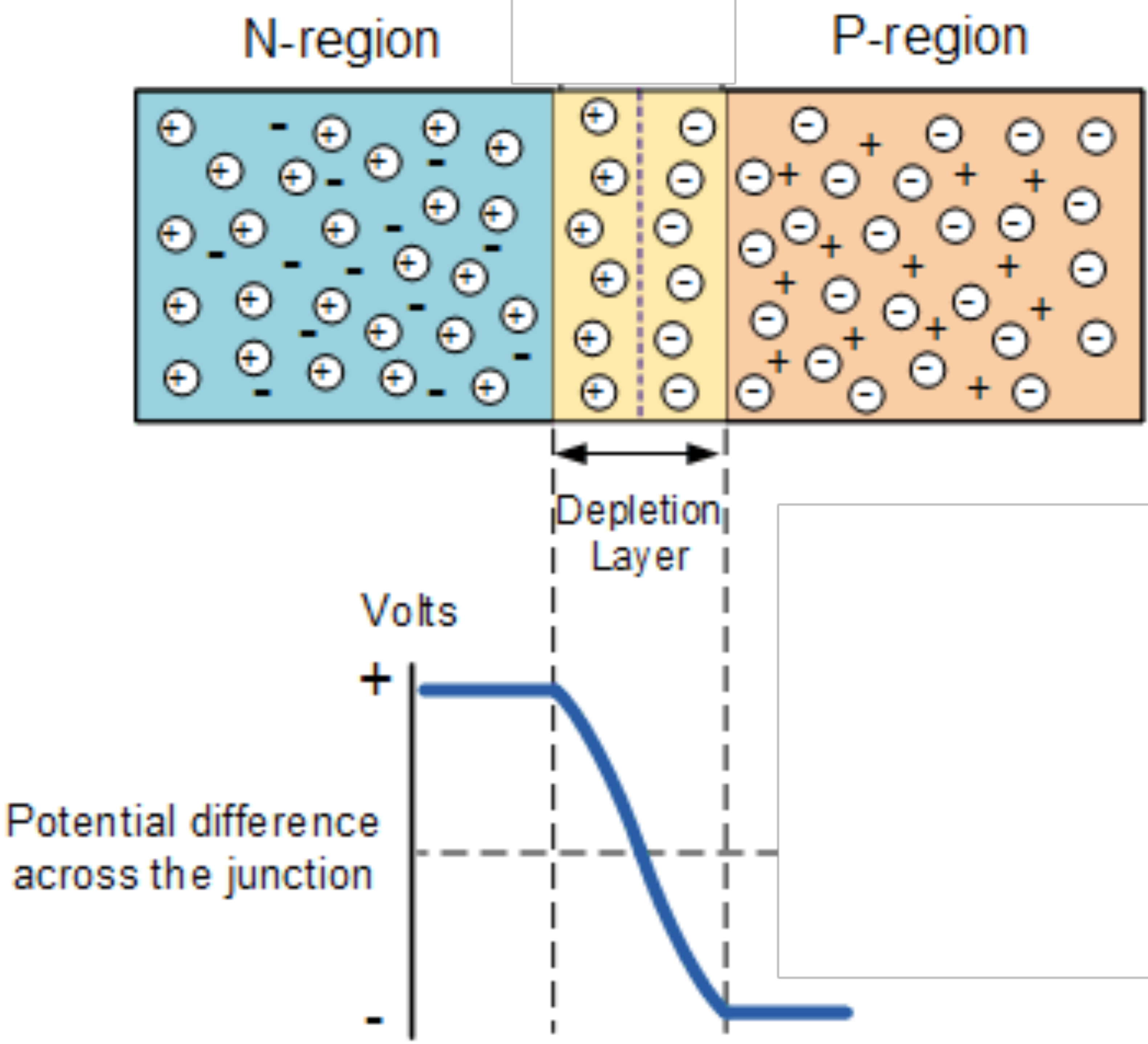


The electrons and holes neutralize each other in a thin <u>depletion layer</u> that has the low conductivity of pure silicon.

The N-region and P-region do conduct.

PN Junction 2

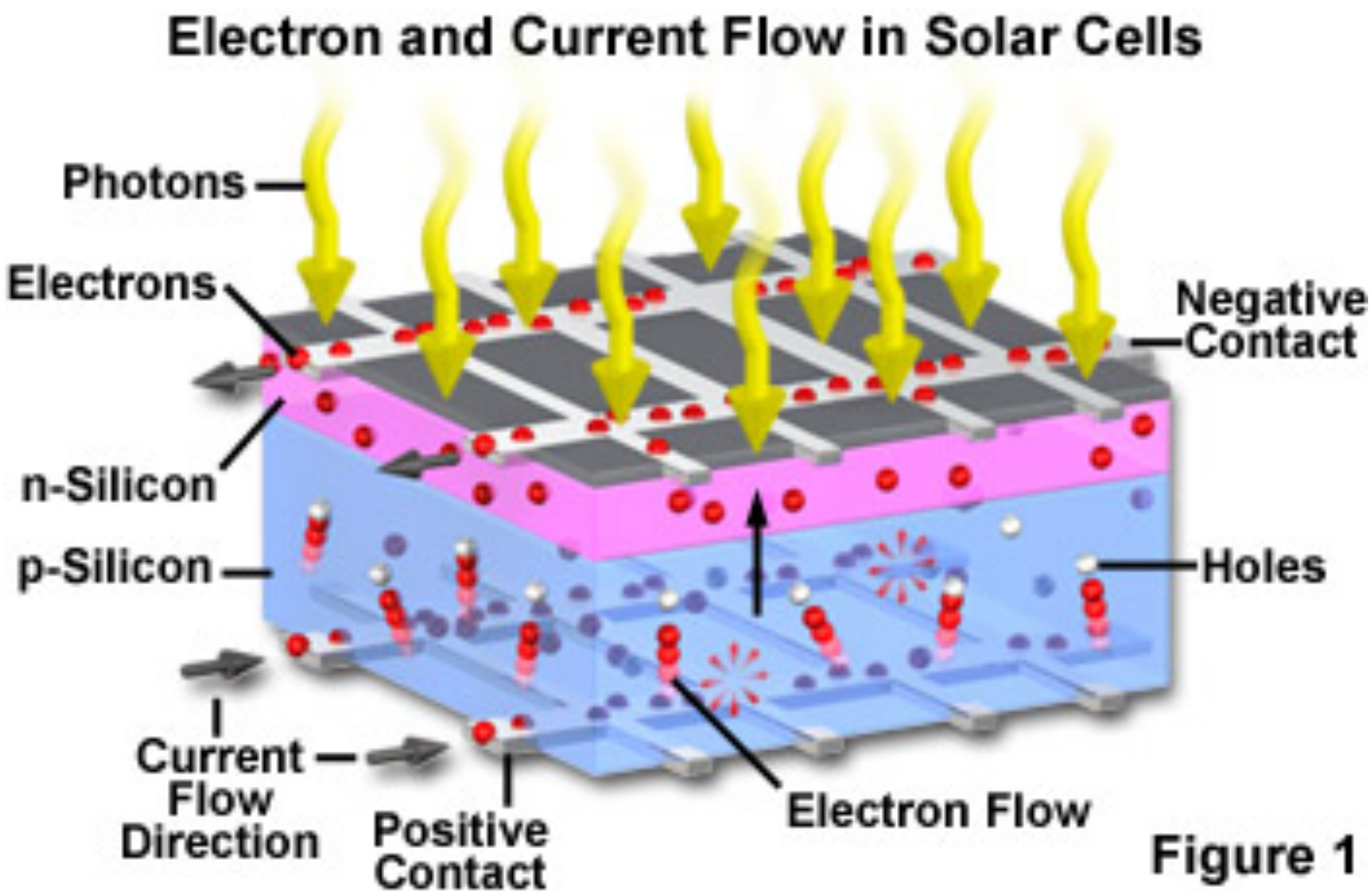
The charge flow to form the depletion layer results in a built-in potential difference across the junction.



Solar Cell

Photons that enter the silicon can detach an electron from an atom, leaving a hole behind.

The built-in electric field separates the electron from the hole, causing an electric current to flow.



Rectifier Diode

The thickness of the depletion layer can be changed by applying an external voltage to the PN junction.

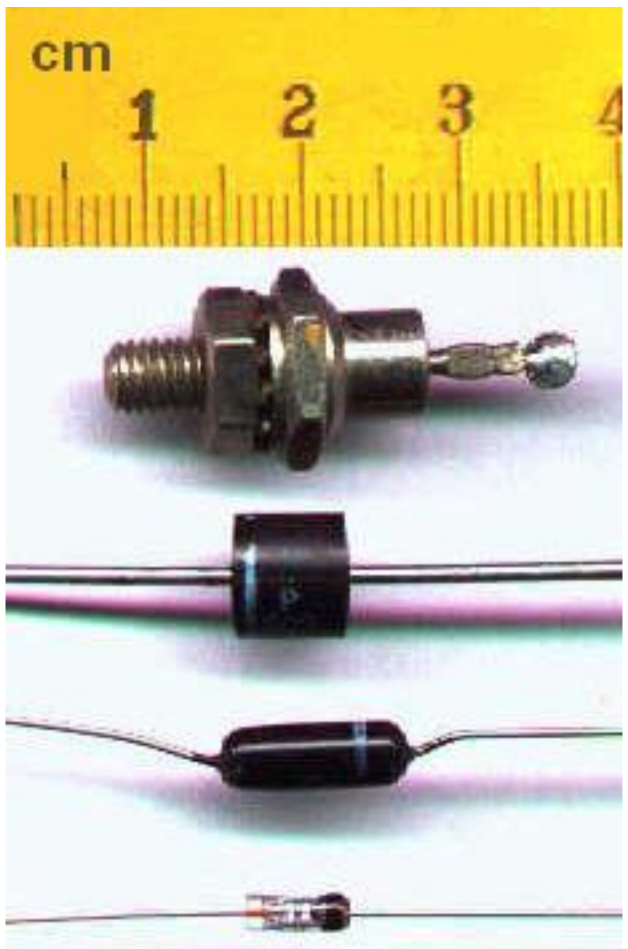
If a large enough positive voltage is applied to the P-side, the depletion layer gets so thin that significant conduction occurs. This is the standard rectifier.

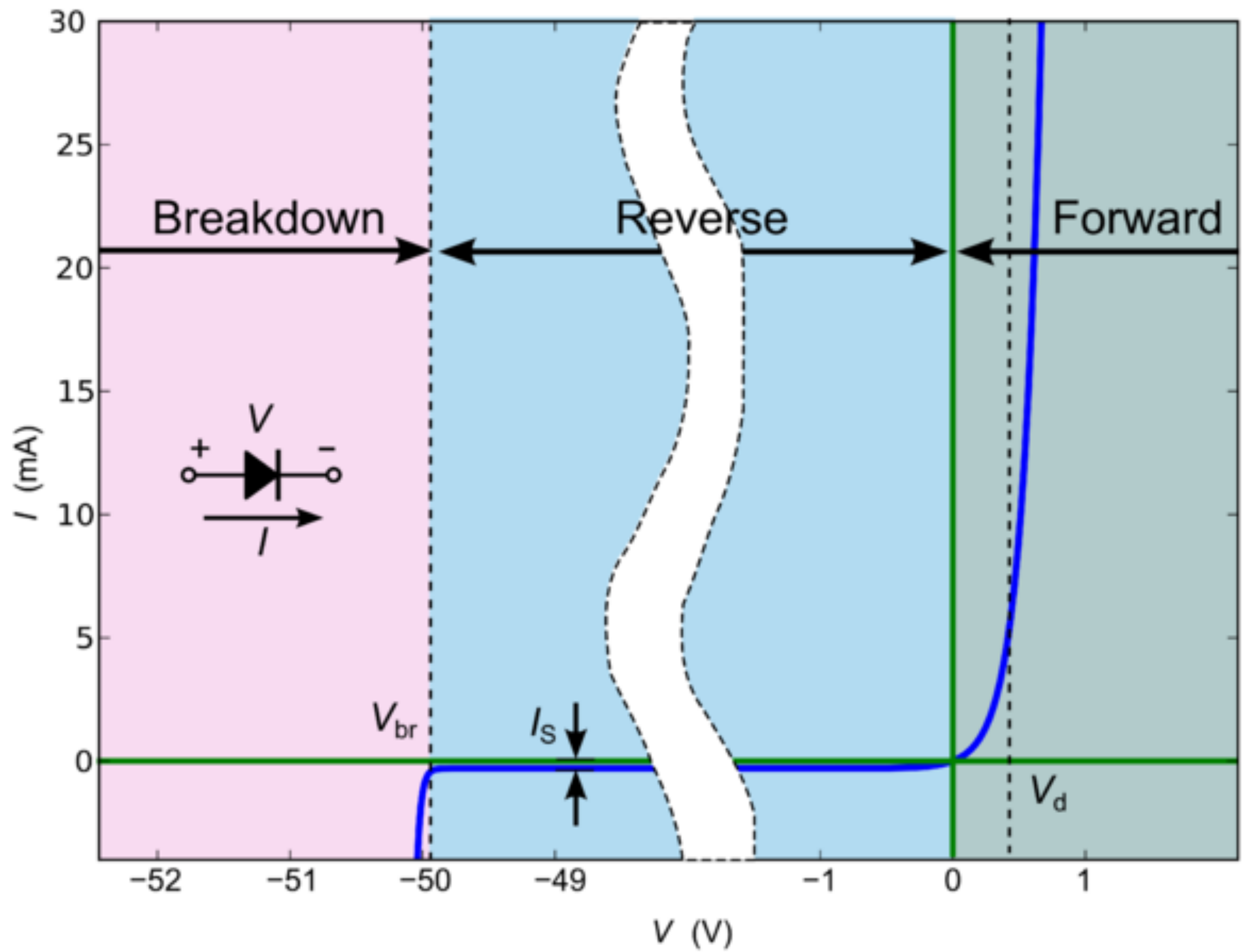
For negative voltage, the depletion layer gets thicker. There is a reverse current that is almost independent of voltage, but usually immeasurably small.

For large enough negative voltage, the few charges that cross the depletion region get enough energy that they can knock electrons from the valence band up to the conduction band, causing a large current. This is reverse breakdown.

Zener diodes are engineered to have a precise reverse breakdown voltage.

Silicon Diodes





Shockley Diode Equation

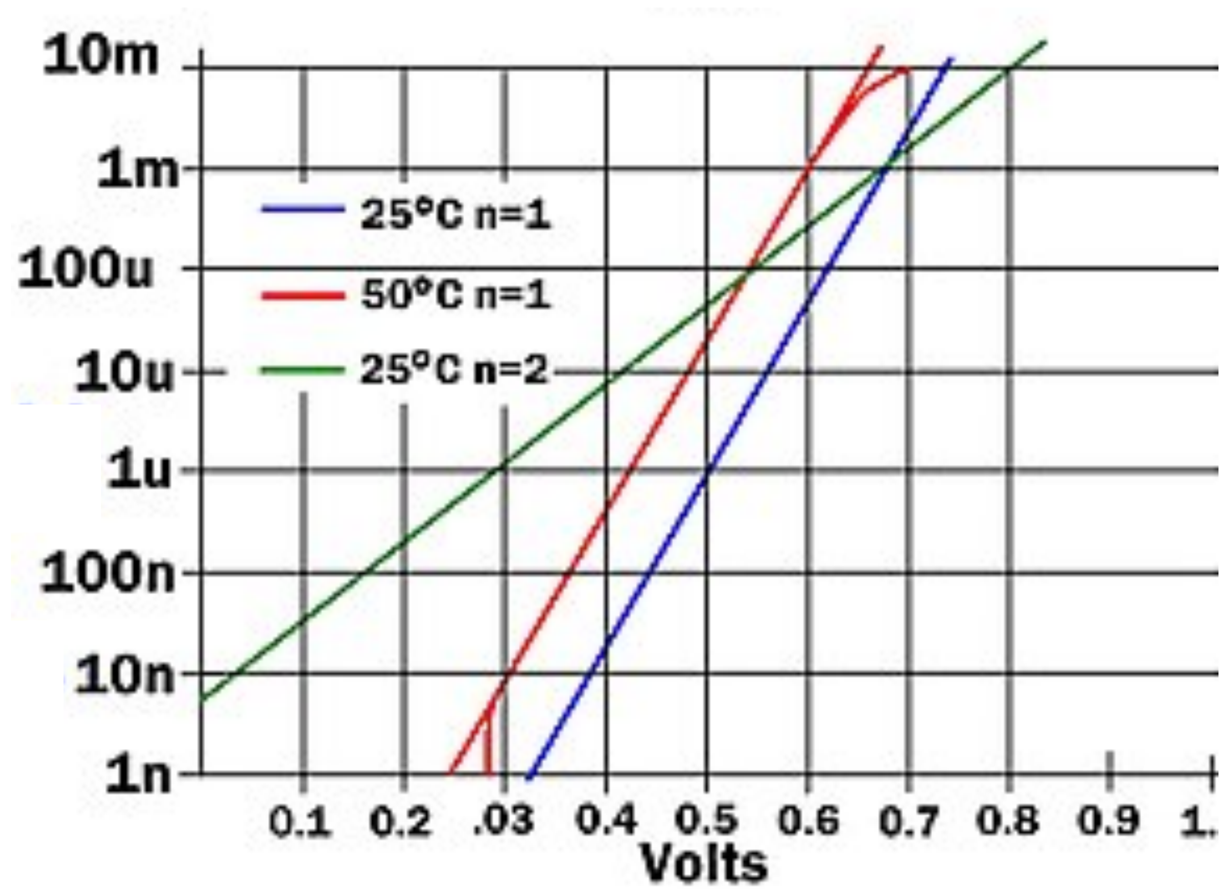
For voltages greater than the reverse breakdown voltage, the current-voltage relationship is described by the Shockley Equation $I(V) = I_S \cdot \left(\exp\frac{qV}{nk_BT} - 1 \right)$

At $V = 0$, the exponential is $+1$, and the current is zero.

I_S is the saturation current, which is the current at negative voltage. It depends on the area of the junction, doping level, etc.

$\frac{k_BT}{q} = V_T$ is called the thermal voltage. It's ~25 milliVolts at room temperature.

The n is called the "ideality factor." It's ~1 for Germanium, ~2 for Silicon.

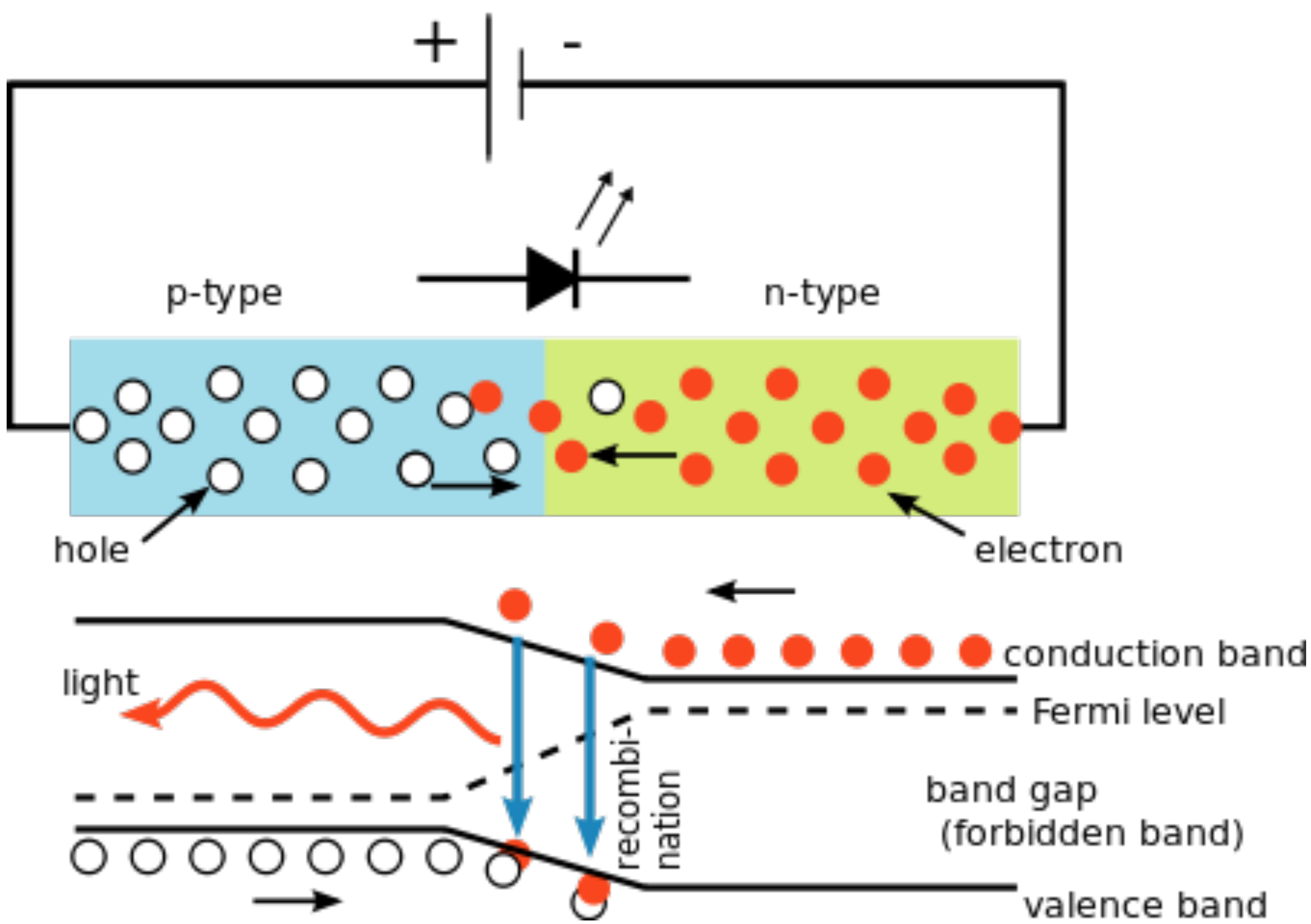


Light Emitting Diode

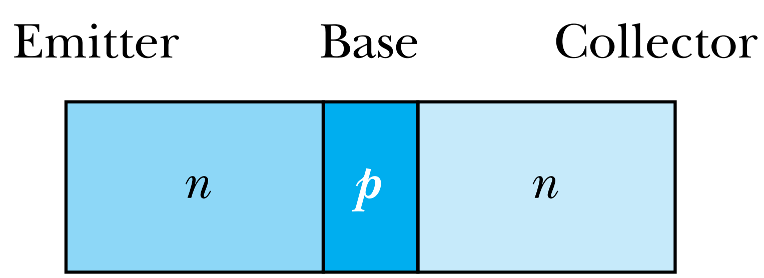
When a diode is conducting, there are lots of electrons and holes in the depletion region.

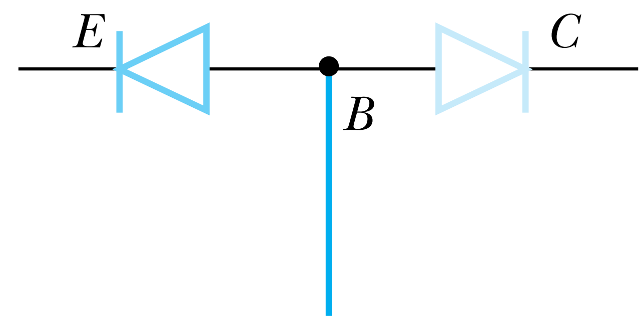
They can re-combine, essentially annihilating each other. That's a nuisance if you are trying to make a good rectifier.

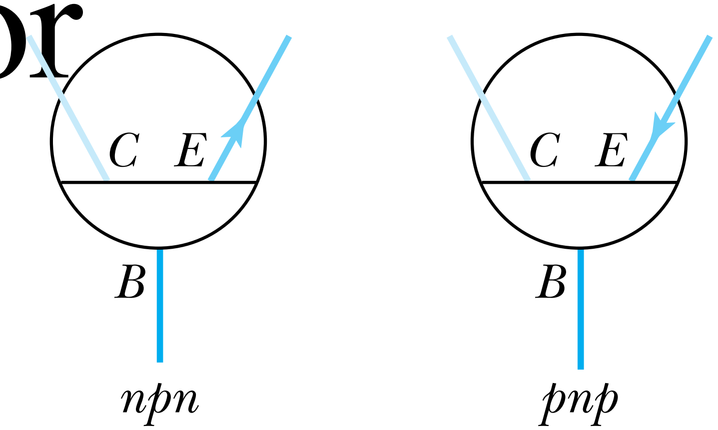
But the energy of the recombination can turn into a photon.



Bipolar Transistor







A bipolar transistor has two PN junctions back to back. It can be NPN or PNP.

The middle region is called the base. It's usually made quite thin.

One end is intended to be forward-biased relative to the base, so its voltage is normally about 0.6 V different from the base. It's called the emitter, because that junction emits either electrons or holes into the base region.

The other end is reverse-biased relative to the base and is called the collector.

The collector current would normally be very small. But the electrons or holes from the emitter-base junction in the depletion region allow current to flow.

The emitter-base current controls the collector current.

Field Effect Transistor

Make 2 N-type islands in P-type silicon. Grow an insulating layer of oxide on top. Etch holes above the N-type islands and deposit metal connecting wires. Deposit another metal wire over the region between the N-type islands, without a hole.

One of the PN junctions is always reverse-biased, so there is normally no current flow between the <u>source</u> and the <u>drain</u>.

But a voltage applied to the gate attracts electrons to the region, allowing current to flow.

Since the gate is insulated, there is negligible gate current, so MOSFETs are far more energy-efficient than bipolar transistors.