Semiconductors

Semiconducting materials have electrical properties that fall between true conductors, (like metals) which are always highly conducting and insulators (like glass or plastic or common ceramics) which always conduct very poorly. The physical property is "conductivity", with symbol σ and units of siemens/meter (S/m).   It indicates how much current will flow when an electric field is applied. Metals have high conductivity ($\sigma > 10^7$ S/m), and insulators have very low ($\sigma < 10^{-12}$ S/m).

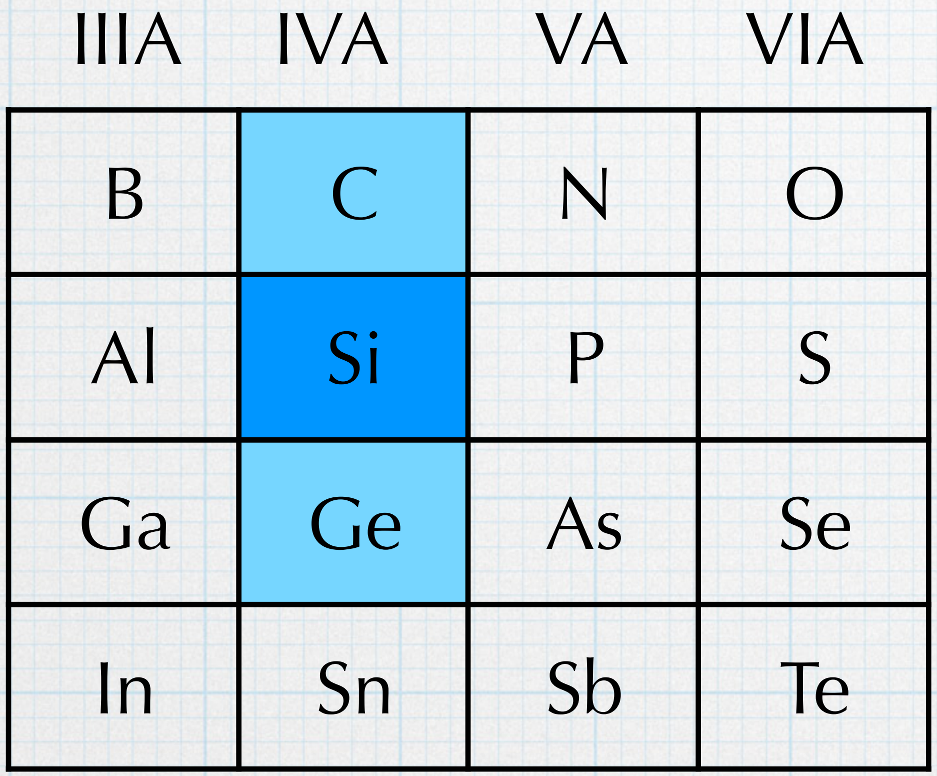
Semiconductors fall between these extremes. More importantly, the conductivity can be modified over a wide range by adding tiny amounts (a few parts per million) of "impurities" to the semiconductor. For example, the conductivity of silicon can range from 10^{-2} S/m to 10^7 S/m, depending on the amount of impurity added. The electric properties of silicon can range from almost metallic to reasonably insulating. In general, conductors and insulators are not as sensitive to compositional changes.

Semiconductor conductivity is also strongly dependent on temperature — increasing the temperature increase the conductivity. In general, conductors and insulators are not as sensitive to temperature changes.

Semiconductors also can collect and generate photons (light), so they are important in hybrid "opto-electronic" applications.

In the semiconductor industry, silicon is the most important material, although it is by no means the only option. Currently, it is the best option for making integrated circuits.

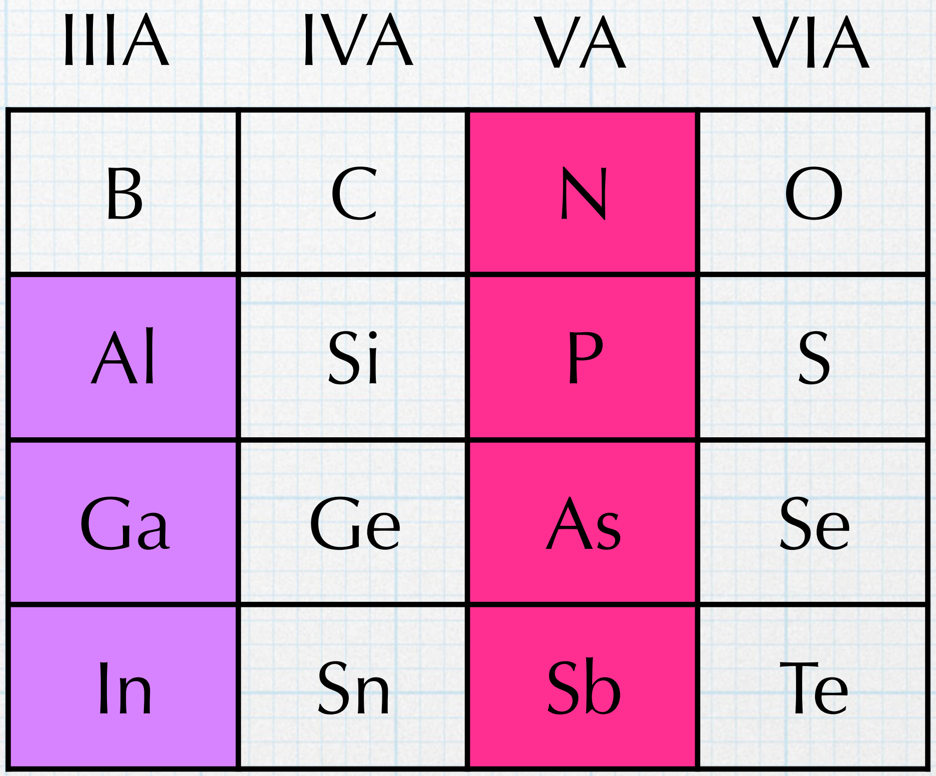
Elemental semiconductors



IIIA	IVA	VA	VIA
B	C	N	O
Al	Si	P	S
Ga	Ge	As	Se
In	Sn	Sb	Te

Si, Ge, C (diamond), SiGe, SiC.

Compound semiconductors



IIIA	IVA	VA	VIA
B	C	N	O
Al	Si	P	S
Ga	Ge	As	Se
In	Sn	Sb	Te

GaAs, AlAs, InP, InAs, GaN, GaSb, (Al,Ga)As, (Ga,In)As, GaIn AsP.

Heterojunction devices — two semiconductors used in a single structure, GaAs/(Al,Ga)As

Some important semiconductors

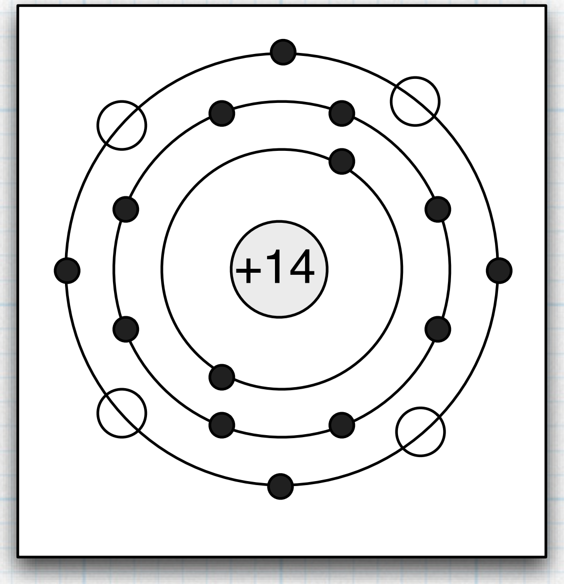
- silicon (Si) — vast majority of integrated circuits, various sensors, light senors, solar cells, miniature mechanical devices
- gallium arsenide (GaAs) — prototypical compound semiconductor, heterojunction bipolar and field-effect transistors for higher performance, infra-red light-emitting diodes (remote controls, etc.)
- gallium indium arsenide (GaInAs) — derivative from GaAs, used in highest speed heterojunction field-effect transistors
- gallium indium arsenide phosphide (Ga,In)(As,P) — used for lasers and detectors in fiber optic systems
- gallium nitride (GaN) — used for high performance microwave and high-power switching transistors and for making ultra-violet LEDs that are the key part of white LEDs for illumination.
- silicon carbide (SiC) — used for high-power switching transistors.

Why is silicon dominant?

- It is abundant — it is the 4th most common element. It is a major component of sand.
- Abundance makes it inexpensive.
- It is relatively easy to purify and crystallize into wafers that are the starting form for integrating circuits.
- The electrical properties can be varied over a wide ranges.
- It is easy to form an insulating layer of silicon dioxide (SiO_2) on the surface of the silicon. The interface between Si and SiO_2 is particularly conducive to making transistors. (This is absolutely key.)
- There are well developed fabrication techniques for turning raw silicon into transistors.
- Integrated circuits can be made that contain billions of transistors whose sizes that are a few billionths of a meter and that switch on and off in a fraction of a billionth of a second.

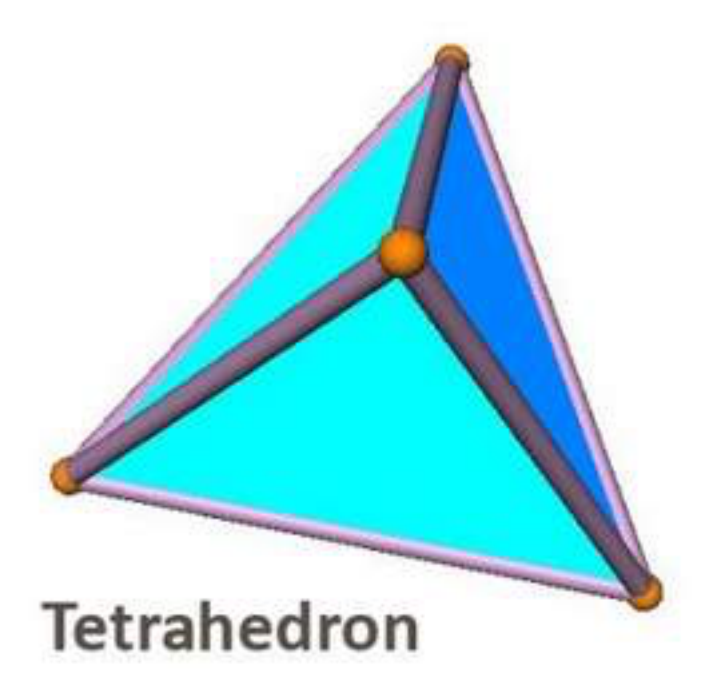
Silicon

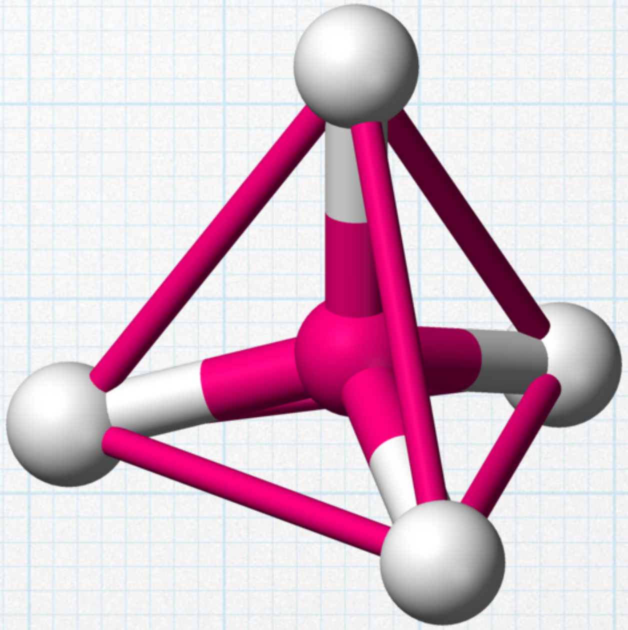
From chemistry: The silicon atom has 14 electrons arranged in the shells of 2, 8, and 4 electrons. The inner two shells are completely filled and play no role in the chemistry of silicon. The outer shell is only half full.

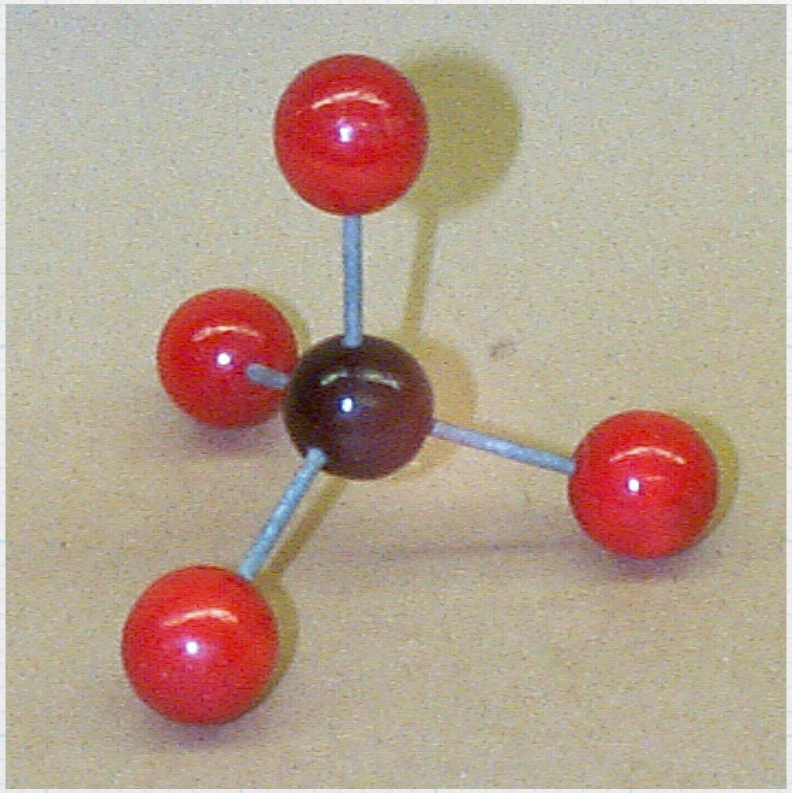


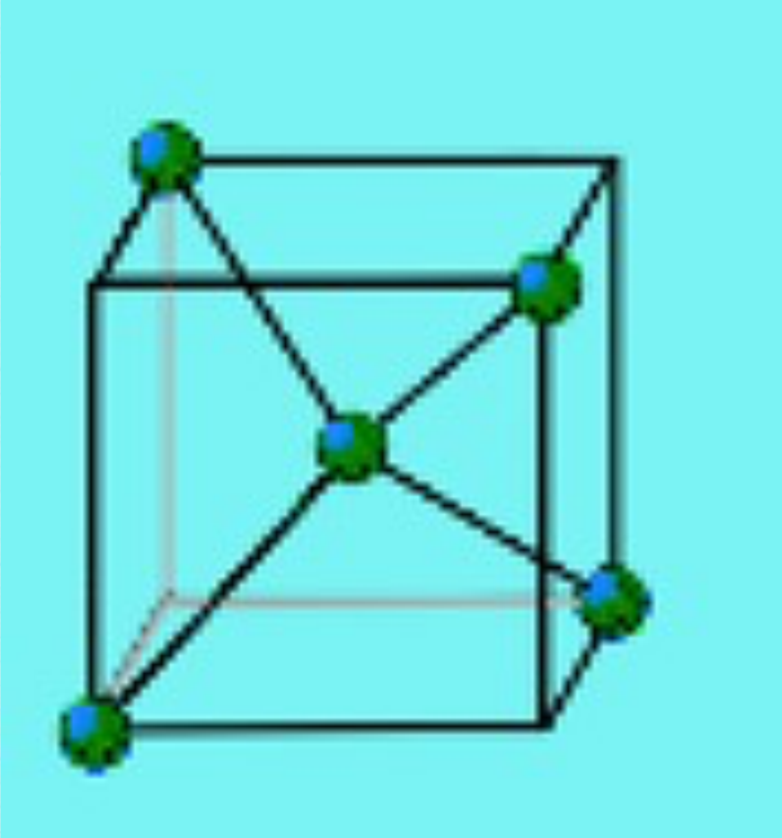
Silicon atoms can bond through the sharing of electrons (covalent bonding). By sharing electrons with 4 other atoms, the atom can "feel like its outer valence shell is full". This arrangement leads to the lowest total energy.

For a silicon atom to most effectively bond with 4 adjacent atoms in three dimensions, the group will arrange into a “tetrahedral” formation.

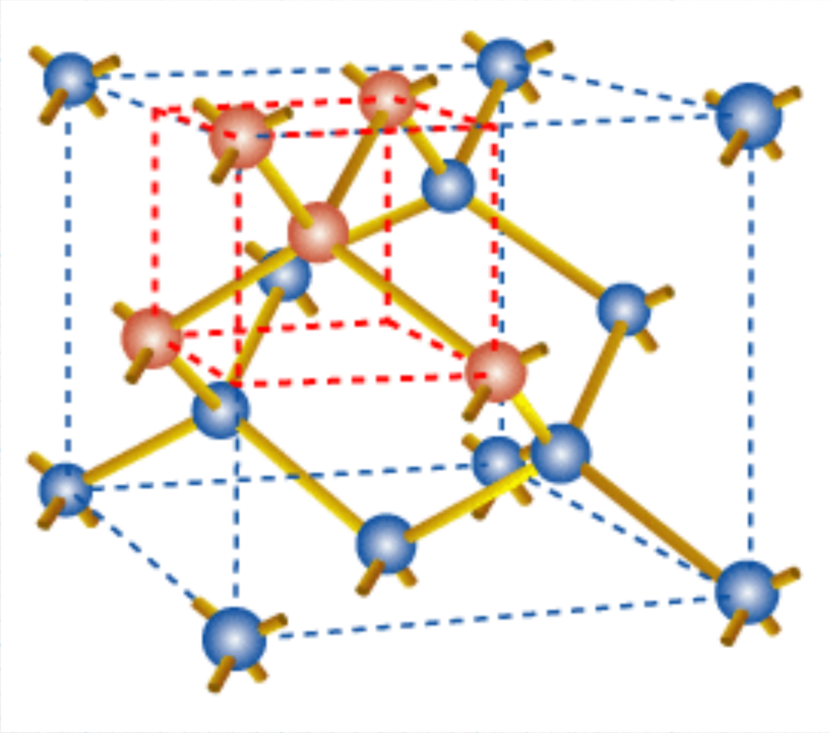


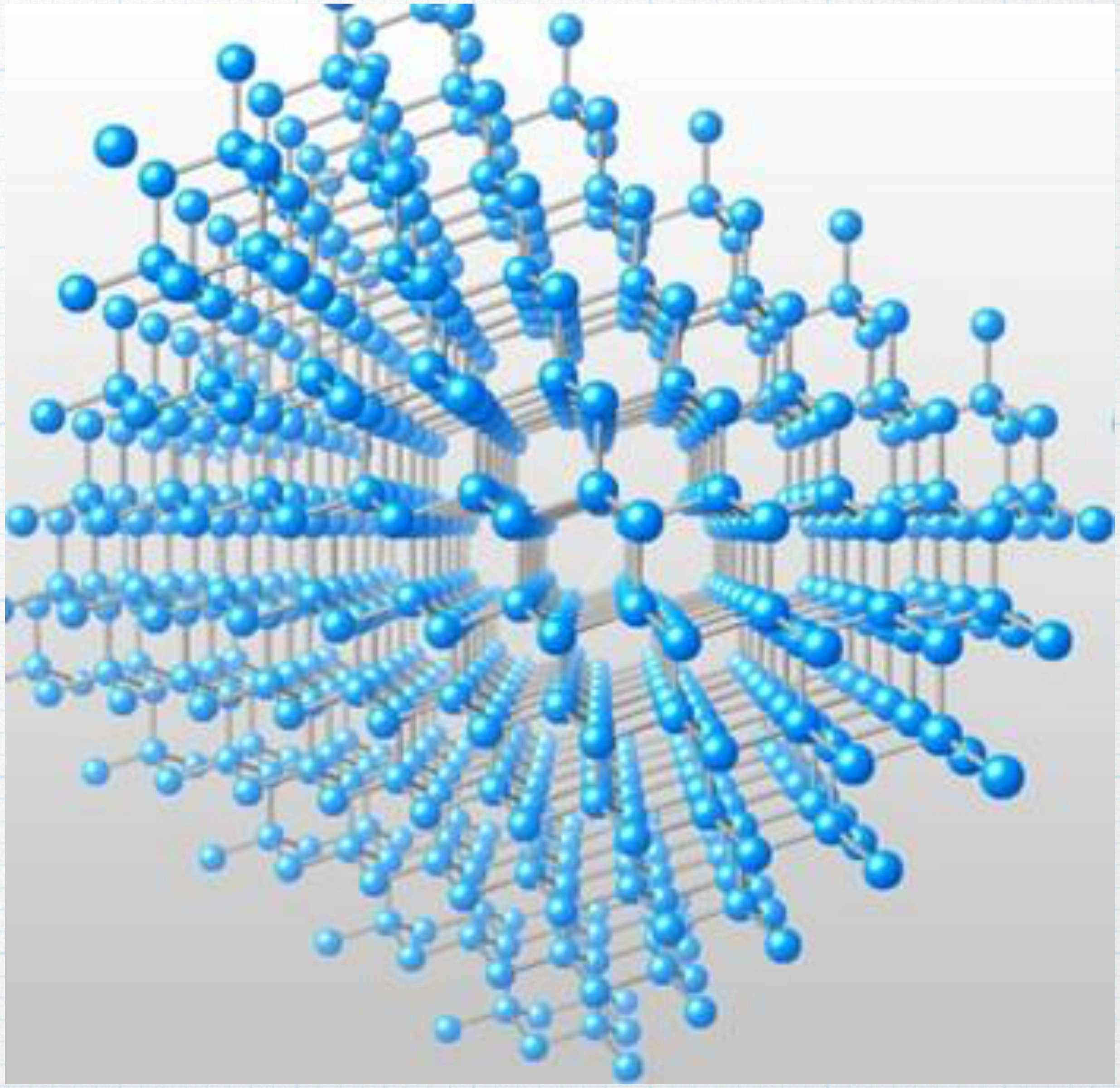






Extending the tetrahedral arrangement to a few quintillions atoms, a silicon crystal is formed. Every atom is in a specific location in the crystal.



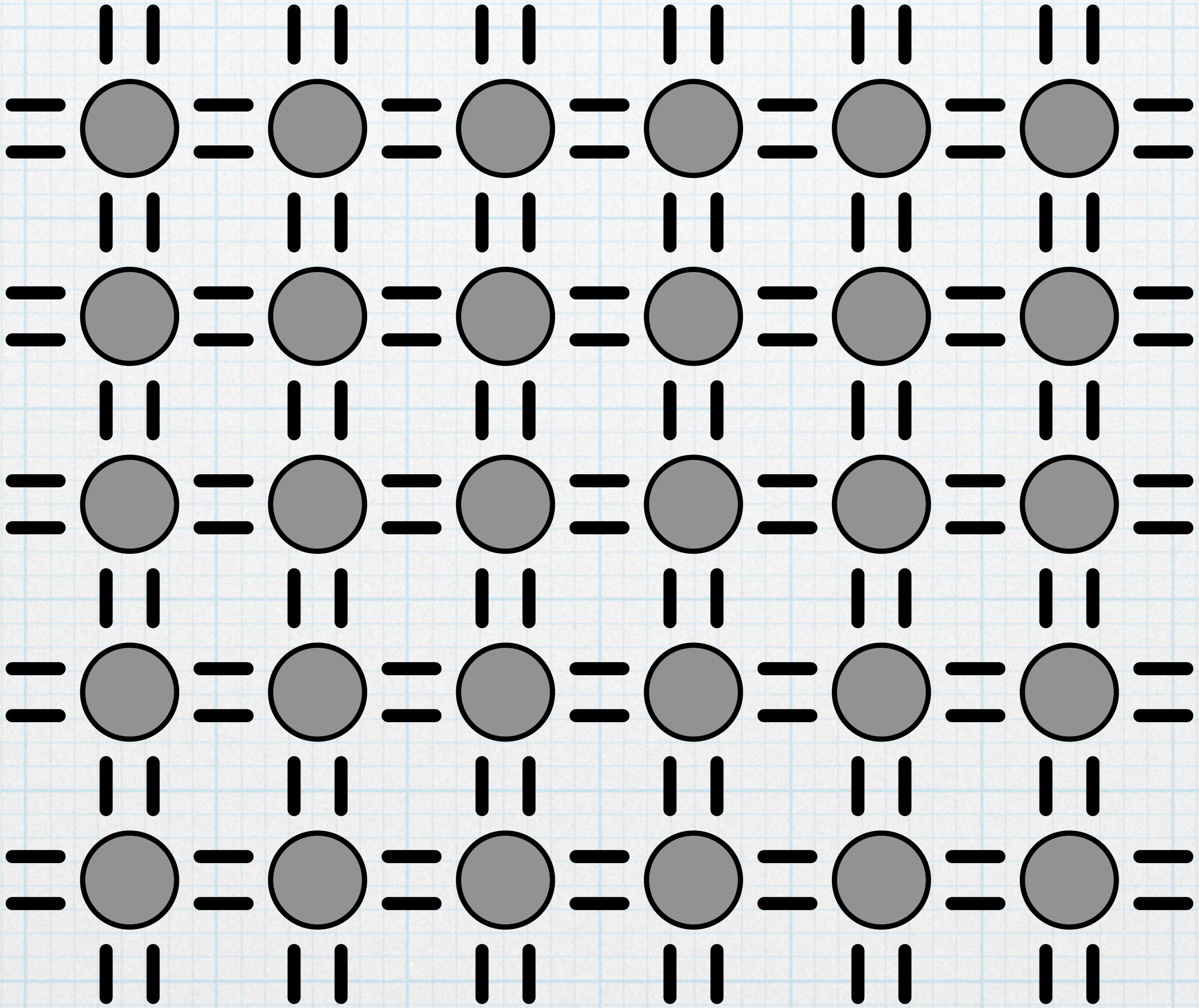


In a silicon crystal, there are about 5×10^{22} atoms in one cubic centimeter.

There are non-crystalline forms of silicon, also.

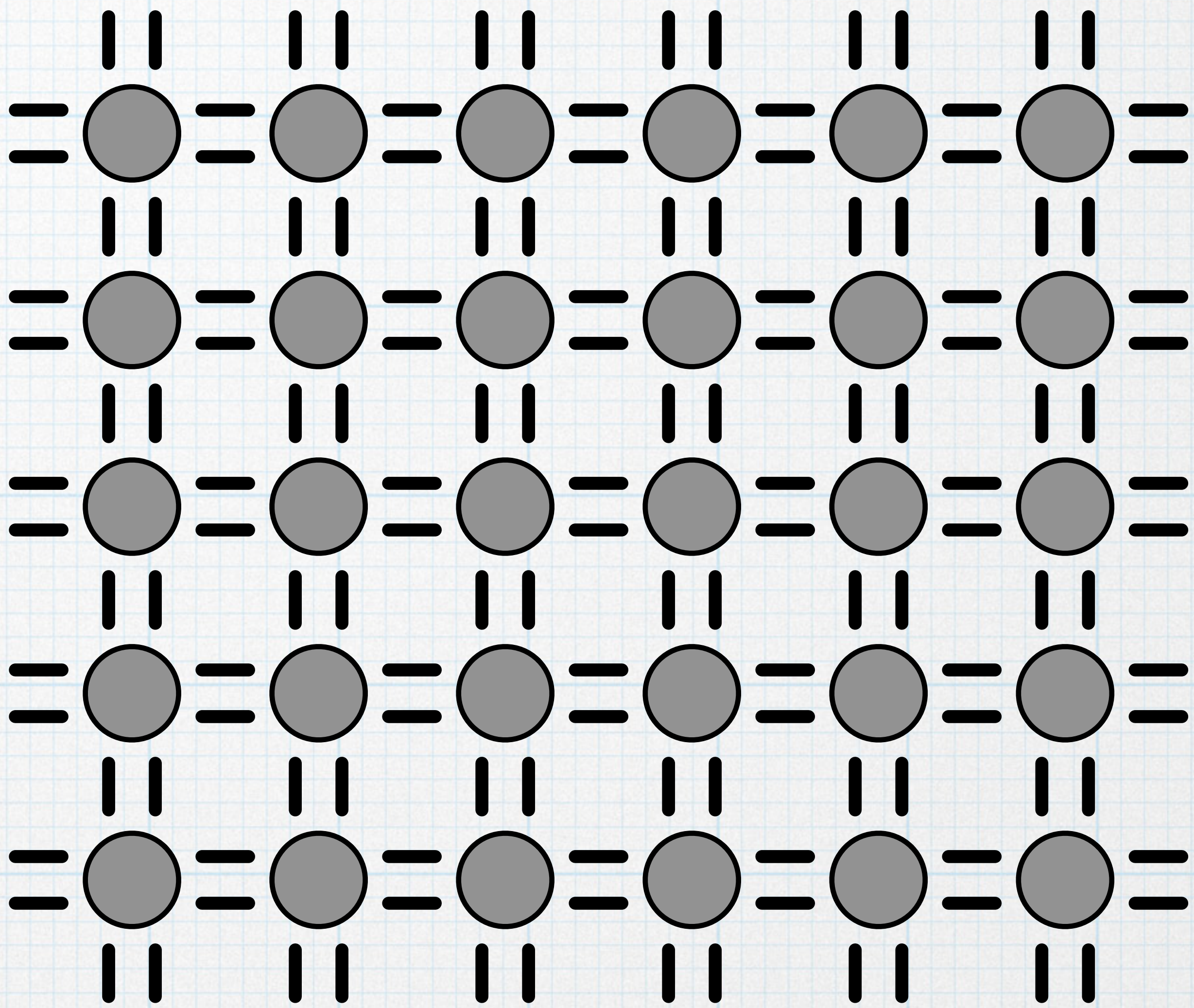
Silicon crystal

It is difficult to depict the 3-D lattice on a flate, but we can get a qualitative idea of what is going on with the 2-D version of crystal, as shown below. The dots represent silicon atoms. The lines connecting dots represent shared bonding electrons. The figure represents a *very* tiny section of a 2-D crystal.



In a perfect silicon crystal, all of the electrons are serving to bond atoms together.

Everything is always in motion

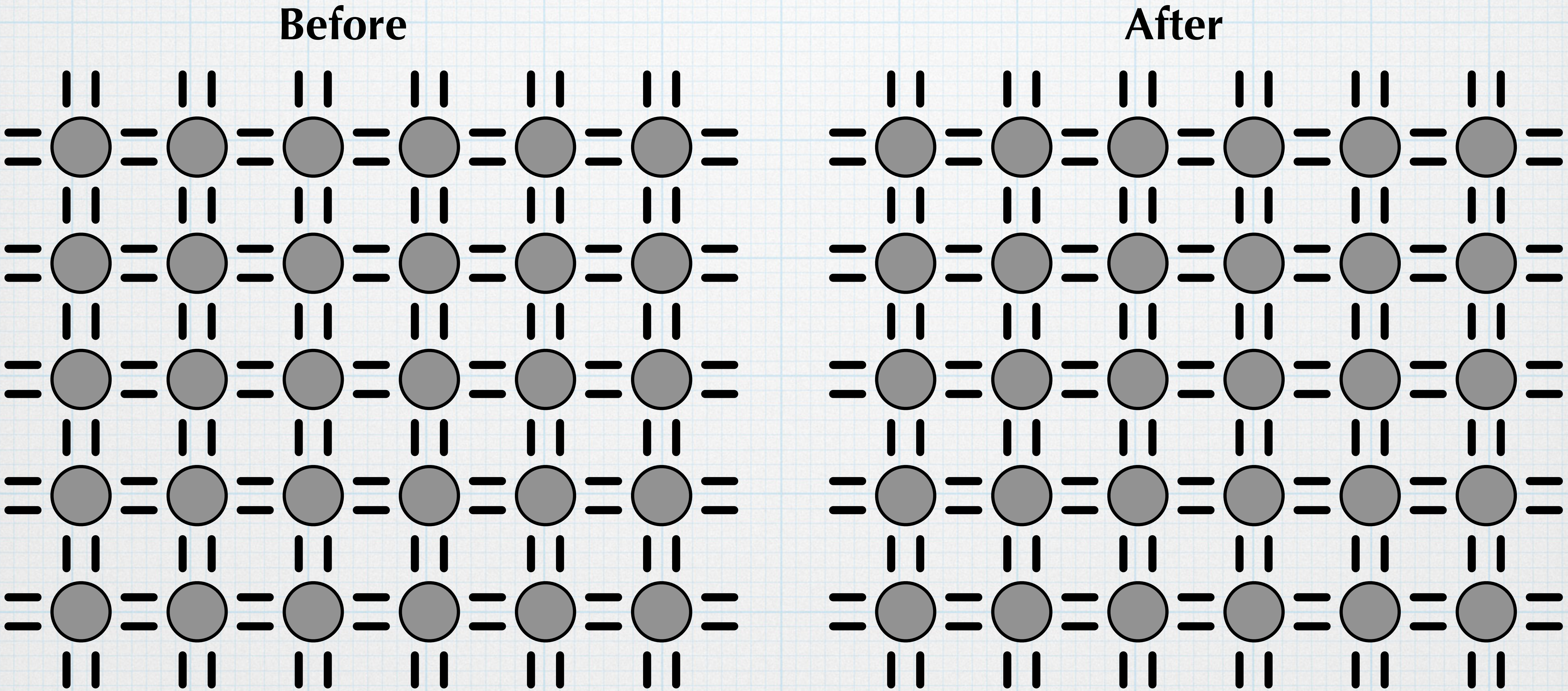


Although the picture implies a static situation, in reality everything is in constant motion. The atoms are vibrating about their lattice sites. The electrons are also vibrating and swapping positions. The electrons are in constant motion — some moving up, some moving down, some moving left, some moving right. All this motion is part of fundamental thermodynamics.

But if we average the motion over time, then static picture above emerge — the vibrations and movements smooth to zero The atoms are properly located at lattice sites with elections between, bonding everything together.

In a perfect crystal, no current flows

A simple though experiment: Take a snapshot of a perfect crystal with all particles in their proper places. Then apply an electric field for short period of time. After the field is turned off, take another snapshot of the crystal.



There is no apparent change — the “after” snapshot looks exactly the same as the “before”. If there is no change in the arrangement of the electrons, then the field had no real effect — no net current flowed during the time the field was present.

Bad musical chairs

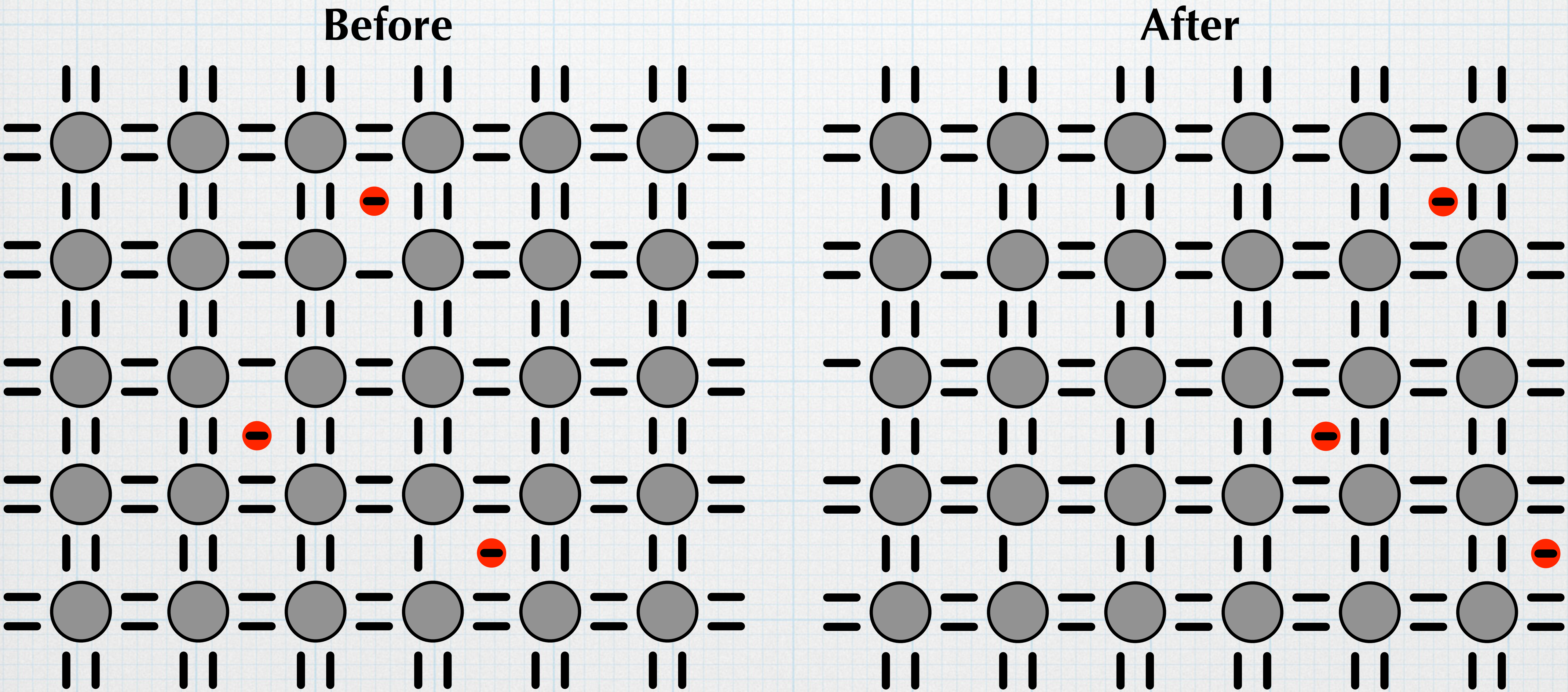
As an analogy, consider a badly implemented game of musical chairs. There are 10 identical children. (Maybe they are clones. Maybe they are all dressed the same, with the same clothes, hats and corona virus masks. How they came to be identical is not the point. The point is that you cannot tell them apart.) There are also 10 chairs.

1. At the start of the game, you photograph the 10 children sitting in the 10 chairs.
2. You close your eyes and the music starts. The children begin marching around the chairs.
3. The music stops and the children scramble to sit down. Of course, with 10 chairs, they all end up with a seat. (That's why this is a bad game of musical chairs.)
4. You open your eyes and take another picture of the children in the chairs.

Comparing the before and after photos, you are unable to see any change. Did anything happen while your eyes were closed? Apparently there was a lot of action while the music played, but there was no net effect overall. So it is with electrons in a perfect crystal. Even if the electrons were moving while field was applied, there was no apparent change at the end. We must interpret this as meaning that no current flowed.

Some electrons break loose

Now suppose that a few rogue electrons have managed to break loose from the bonds and have become "free". If we repeat the electric field experiment again with a field that pushes electrons to the right, we note a distinct difference between the "before" and "after" pictures. The free electrons have moved to the right. There is a measurable change in the electron configuration — a current must have flowed!



Our conclusion is that in order to have currents in the crystal, it cannot be "perfect", there must be some free electrons available.

Thermodynamics and free electrons

How can electrons break loose? Recall that in a little cube of silicon, 1 mm on side, there are about $5\text{x}10^{19}$ (50 quintillion) atoms and so four times that, or $2\text{x}10^{20}$ (200 quintillion) valence electrons. The electrons are all wiggling like crazy, but mostly they stay between the atoms in the bond positions.

However, according to the rules of thermodynamics, the electrons have a distribution of energies and the number of electrons at a given energy is specified by a probability function — in classical physics it is the Boltzmann function, $f(E) = \exp\left(-\frac{E}{kT}\right)$, where E is the energy, kT is the thermal energy, and $f(E)$ is the probability that an electron has that energy. (In a quantum system like a semiconductor crystal, the probability function is different, but the basic concepts are the same. We will not delve into the differences here — we save it for EE 332.)

The probability decreases exponentially as the energy increases, but *it never goes to zero*. There is a small, but finite, probability that some electrons will have very high energies — high enough that they can break out of the bonds and become free to move through the crystal. Because there are so many electrons, it is certain that *some* will have broken loose. The fraction of free electrons may be small, but it will not be zero.

The band gap

For a given semiconductor, the energy required for an electron to break loose is known as the *band-gap energy,* denoted as E_G. Each semiconductor has a distinctive E_G. The band gap can be viewed as measure of how strongly the electrons are held in the bonds. The bigger the bandgap, the harder is it for electrons to break loose. Thus, the bandgap energy becomes a key indicator of the electricalcapabilities of the material. In fact, we can re-visit our earlier characterization of electronic materials using the idea of a band gap.

Insulators have a very big band gaps. Even though there are many electrons bonding the atoms together, the energy required to break loose is so big that virtually none of them can do it. There can never be free electrons and the insulator has virtually no current — the conductivity is extremely small.

Conductors — most metals — have very small band gaps. It is very easy to for electrons to break and there will always be lots of electrons present to carry current — the conductivity is always very big.

Semiconductors fall into between. The gap is not too big — depending on the temperature, there will be some numbers electrons that have broken loose and can carry current. The conductivity will have an intermediate value.

Bad musical chairs with a rogue student

Returning to our musical chairs analogy, suppose that one of the children is not interested in musical chairs. Instead of marching around the chairs with the rest of the group, they wander over the corner of the music room where the percussion instruments are set up.

1. At the start of the game, you photograph the 10 children, 9 sitting in chairs and one over banging on a snare drum.
2. You close your eyes and the music starts. The 9 children dutifully begin circling the chairs. The "rogue" child moves over the xylophone.
3. The music stops and the 9 children scramble to sit down. Of course, there are plenty of chairs. Meanwhile, the 10th child has moved to the cymbals and is clanging them together.
4. You open your eyes and take another picture of the children in the chairs and the rogue youngster, who is, by now, thumping on the bass drum.

Now in comparing the photos, you see that the "free" child has clearly been moving around while your eyes were closed, and you can make an inference about their between-snapshot movements, which could be described as a "current".

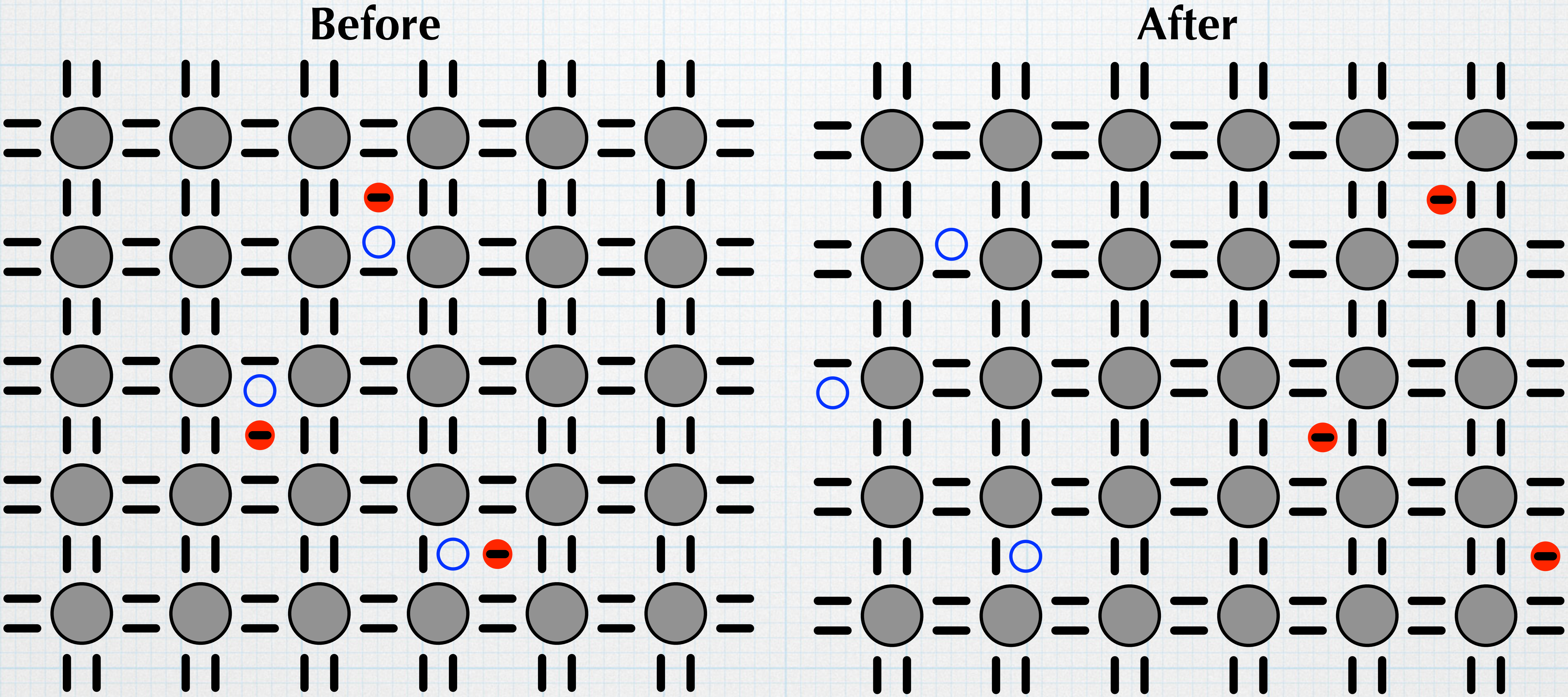
Holes

So free electrons in the crystal are a means by which current can flow. If the band gap energy of the material is not too big and the temperature is not too low, there should be some free electrons available. But in focusing on the free electrons (or the free child in the musical chairs game), we are overlooking another aspect — the "empty spot" vacated when the electron left the bond position (or the child left the game) is also moving.

Consider the musical chairs game again. While we are tracking the "free" child as they move around in another part of the room, the other children are still playing the game. During each round the, 9 kids march around the 10 chairs, and when then music stops they sit down. At the end of each round, there will be an empty chair. From one round to the next, the location of the empty chair will be in a different spot, depending on the random positions of the children when the music stopped. We can clearly see the changing position of the empty chair each time, and that change also represents current flow.

In our musical chairs scenario, we see that there are two mechanisms leading to changes in the positions of the kids — the motion of the free kid goofing around with percussion instruments and the motion of the game-playing kids, which we track by watching the changing location of the empty chair.

It is much the same for electrons in the crystal. Look again at the free electrons in the "before" figure, where we have indicated the empty spots — holes — where the free electrons were previously bound. The applied electric field pushes electrons to the right.

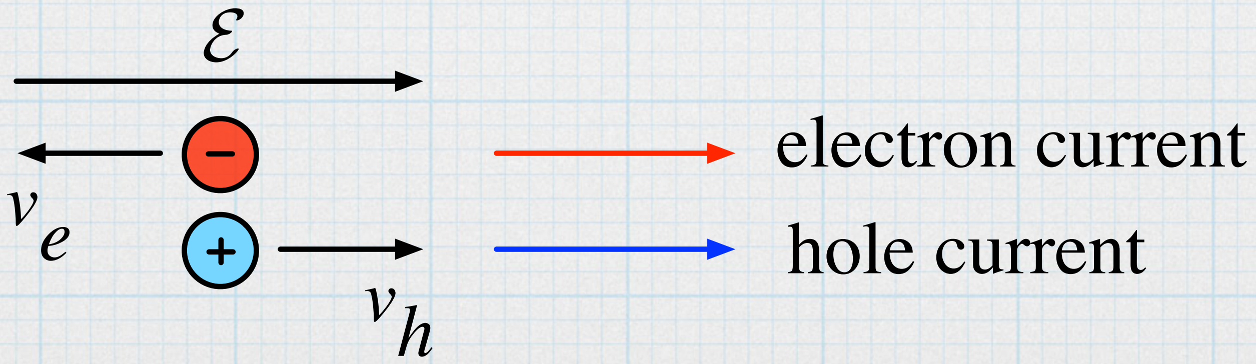


The free electrons move to the right. In addition, we now see that the bonding electron can also hop to the right to fill an empty. For each of the original empty spots, there have been electron hops to the right, which, causes the empty spot to move to the left. The empty spots move as if they as if they were electrons, but with positive charge.

The change in the arrangement of the bonding electrons also represents current flow. We can keep track of it by monitoring the movement of the empty spots. In some sense, this is an accounting trick. The current is due to the movement of the electrons in the bond positions, but it is easier to keep track of a few empty spots that it is to keep track of a 100 quintillion or so electrons — just like it is easier to watch the movement of the empty chair in the game that it is to keep track of 9 identical children.

So we see that current in a semiconductor can be carried by free electrons or "holes". When applying an electric field to push carriers around, the electrons move "normally" — opposite the direction of the electric field — and the holes move as if they were positively charged — in the same direction as the field.

However, recall that electrons moving to the left represent *negative* charge moving to the left. This has the same effect as positive charge moving to the right. So while an electric field causes electrons and holes to move in opposite directions, the two current are in the same direction.



Concentrations

Because the numbers of electrons and holes are so very big, we describe the quantities in terms of concentrations. The concentration of free electrons and holes is the number per unit volume.

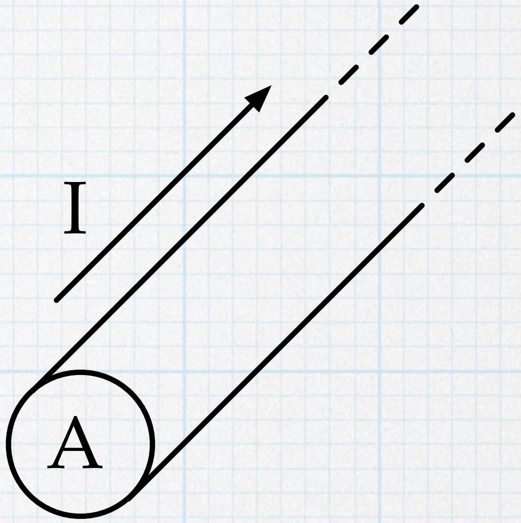
To get an idea of just how many particles there are, we start with the concentration of atoms in a silicon crystal. This number works out to about 5×10^{19} per cubic millimeter. [50 nm^{-3} or 5×10^{22} cm^{-3} or 5×10^{28} m^{-3}.]

There are 14 electrons with each atom, but only 4 are in the outer valence shell and hence available to break loose. So there are about 2×10^{23} cm^{-3} electrons that might possibly be available to carry current. Usually, only a very tiny fraction of electrons are "free" — probably less than one in a trillion have sufficient energy to break the bonds. This means that the concentration of free electrons might be something like 10^{10} cm^{-3}. Of course, there would be similar concentration of holes.

$n \rightarrow$ electron concentration

$p \rightarrow$ hole concentration

Current density



In similar fashion, we normalize currents by using the idea of *current density*. If we consider a wire (or any path that constrains the flow of current) with cross-sectional area of A, then the current density is defined as $J = I/A$.
The units are amps per unit area: A/cm^2.

As an example, consider a 14-gauge copper wire (diameter of 1.63 mm), like the wiring in your house. If there is one amp of current flowing, the current density is:

$$J = \frac{I}{\pi r^2} = \frac{1\ \mathrm{A}}{\pi\,(0.0815\ \mathrm{cm})^2} = 48\ \mathrm{A/cm^2}$$

In a trace of an integrated circuit, a wire may have only 1 mA flowing but the rectangular cross-sectional area might be very small, say 5 µm x 0.5 µm. In that case the current density might be:

$$J = \frac{0.001\ \mathrm{A}}{(5\times10^{-4}\ \mathrm{cm})\,(5\times10^{-5}\ \mathrm{cm})} = 40,000\ \mathrm{A/cm^2}$$

Current

Of course, current is moving charge. So there must be some charge (as indicated by the concentration) and there must movement, as given the velocity. And we must have both things — charge with no velocity is zero current as is velocity without charge. So if we consider electrons, it is the product, $q \cdot n \cdot v_n$ that matters. (Each electron carries 1 q of charge and v_n is velocity of the electrons.) If we examine the units of this quantity, we realize that this is current density.

$$J_n = q \cdot n \cdot v_n$$

If we are considering holes, the current density is:

$$J_p = q \cdot p \cdot v_p$$

If both electrons and holes are present then

$$J_{total} = J_n + J_p = q \cdot n \cdot v_n + q \cdot p \cdot v_p$$

Intrinsic semiconductor

However, the situation we have considered so far is a special case, because the numbers of holes and electrons are always identical — the creation of a free electron automatically introduces a hole. We call these electron-hole pairs. This is the inherent result that must occur when we have a perfect crystal in which every atom is silicon, so that the number electrons exactly matches the number of available bonds. In this situation, with the number of electrons exactly matching the number of holes, $n = p$, we have an "intrinsic" semiconductor.

Technologically, intrinsic semiconductors are not very interesting, because the only means for changing carrier concentration is to change the temperature. And then we change both the electrons and holes in identical fashion.

Extrinsic semiconductor (doping)

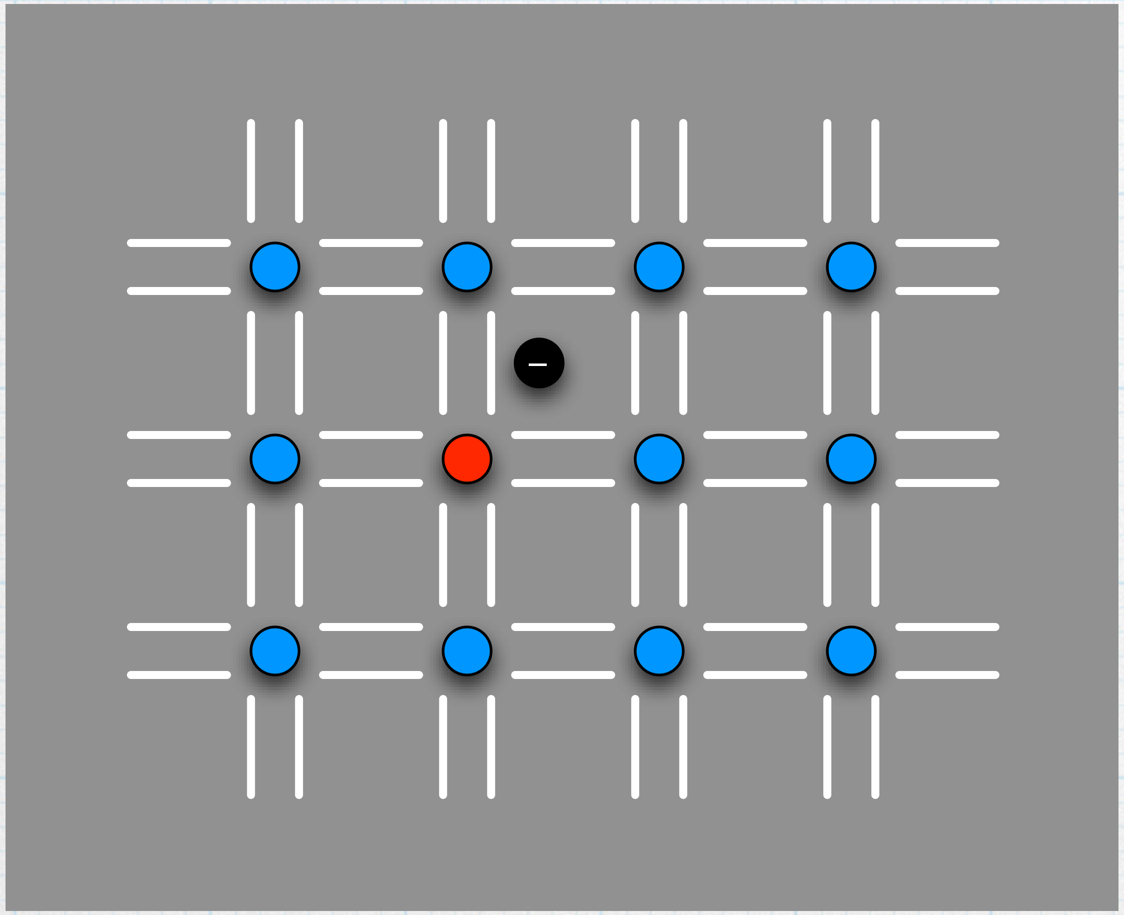
If we want to control the electron and hole concentrations independently, we must make the crystal imperfect in some fashion. In particular, we will want to form defects by controllably adding atoms of other elements. The process of adding impurities is called doping and the added impurities are dopants.

To see the effect of a dopant, consider substituting phosphorus atoms for some of the silicon atoms in the crystal.

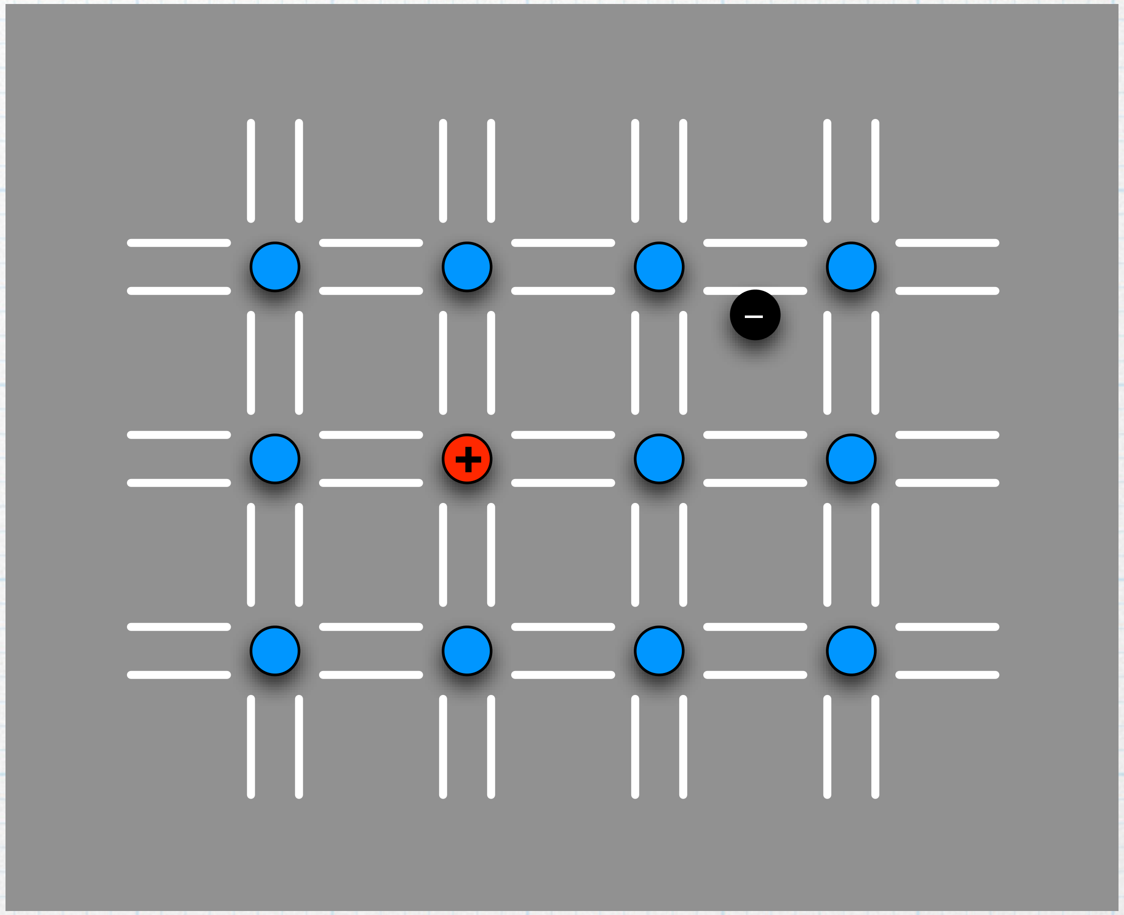
Phosphorus comes from the fifth column of the periodic table. It is very similar to silicon, but has an outer shell with 5 valence electrons.

IIIA	IVA	VA	VIA
B	C	N	O
Al	Si	P	S
Ga	Ge	As	Se
In	Sn	Sb	Te

If a phosphorus atom is substituted for a silicon atom in the crystal, it fits in relatively well — using four electrons to make bonds with the four neighboring silicon atoms. However the fifth electron has no bonding site in which to fit.



The dangling electron will be weakly bound to the phosphorus atom.



It is relatively easy (i.e. requires little energy) for the electron to break free. The phosphorus atom will then carry a charge of +*q*.

Since the added phosphorus atom “donated” one electron to the semiconductor, it is known as a “donor”. (Any element that adds electrons is a donor.) The use of donors gives us a means to control the electron concentration, independently of the holes. It seems intuitively obvious that each added donor atom will add an electron. So we might write:

$$n = N_D,$$

the electron concentration is equal to the donor concentration (if no other impurities are present). The simple relationship is generally correct. So a semiconductor in which donors have been added to increase the concentration of electrons is known as an “*n*-type” semiconductor.

Phosphorus is probably the most commonly used *n*-type dopant, although arsenic is used in some special situations.

Some typical numbers

Concentration of silicon atoms in a silicon crystal:
$5\text{x}10^{28}\ \text{m}^{-3} = 5\text{x}10^{22}\ \text{cm}^{-3} = 5\text{x}10^{19}\ \text{mm}^{-3}$.

Intrinsic electron (and hole) concentration at room temperature:
$n_i = 6\text{x}10^{10}\ \text{cm}^{-3}$ ($\approx 10^{10}\ \text{cm}^{-3}$ or $10^{7}\ \text{mm}^{-3}$.)

Doping concentrations, either p-type or n-type:

N_D (or N_A) ranging from $10^{14}\ \text{cm}^{-3}$ to $10^{21}\ \text{cm}^{-3}$ (2 ppb - 2%)

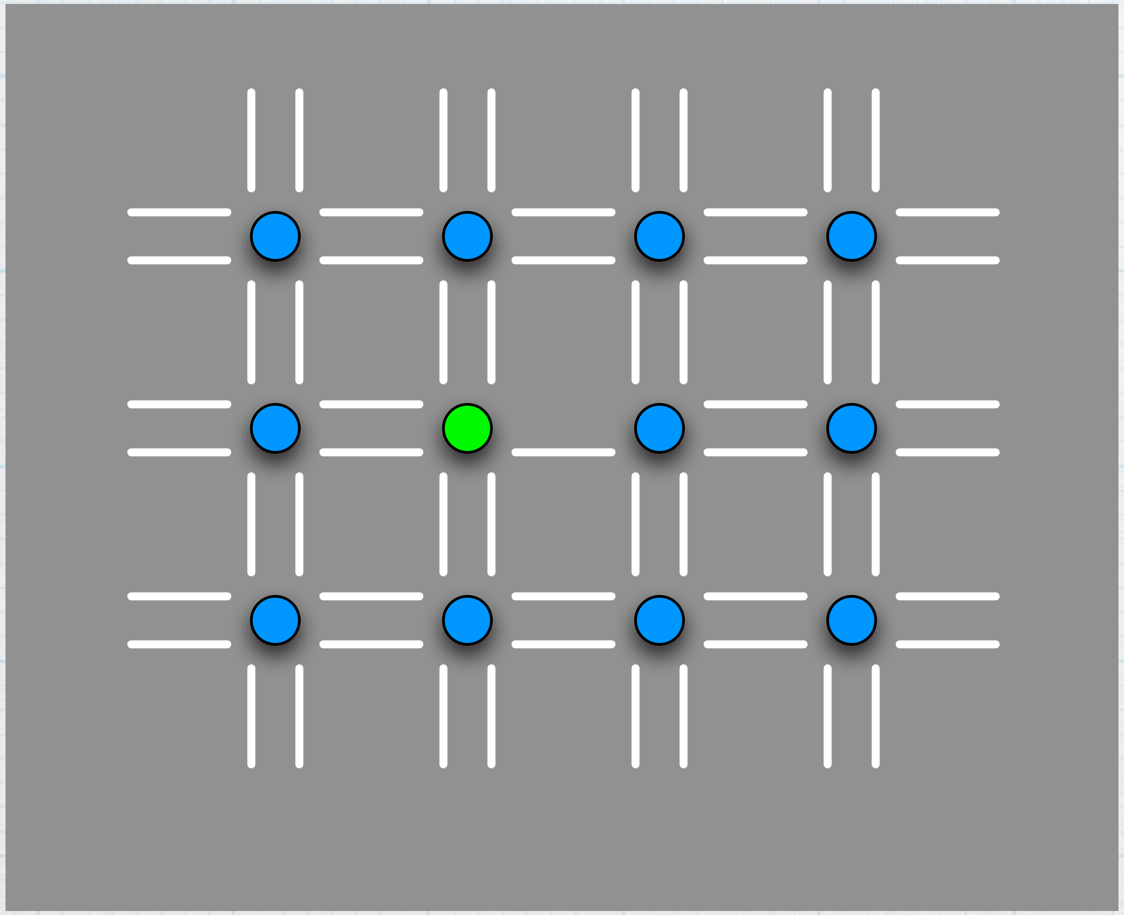
Then, under typical conditions in *n*-doped silicon, the electron concentration would then be $n = N_D$. In *p*-doped silicons the hole concentration would be $p = N_A$.

Acceptors

In an analogous way, holes can be added to the semiconductor.

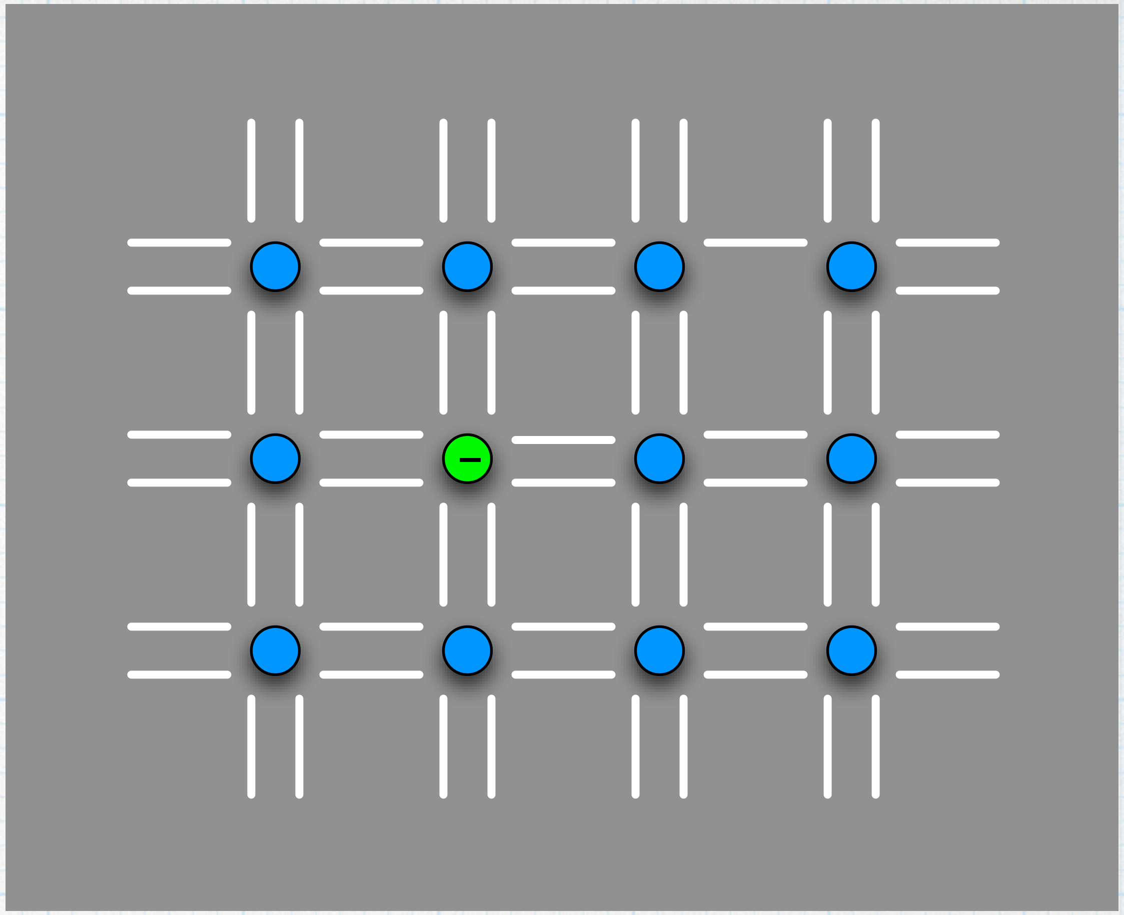
Consider substituting boron atoms for some of the silicon in the lattice. Boron has three outer valence electrons – one short of matching up with the four surrounding silicon atoms.

IIIA	IVA	VA	VIA
B	C	N	O
Al	Si	P	S
Ga	Ge	As	Se
In	Sn	Sb	Te



Each added boron atom introduces a hole into the bonding arrangement. Since the empty spot can accept an electron, boron is known as an acceptor dopant.

Once the hole "breaks loose", the acceptor atom, which is locked in place, will take a net negative charge of -q.



Again, since each acceptor atom adds one hole, we can surmise that

$$p = N_A,$$

the hole concentration is equal to the added acceptor impurity concentration (if no other impurities are present). A semiconductor that has a hole concentration that is greater than the electron concentration is known being "p-type".

The ability to controllably change the doping in different regions of a semiconductor is the basis for building virtually all of the electronic devices that are now in use.

Dopant compensation

Typically, a region of semiconductor will have both types of dopant present. In that case, the dopant with the higher concentration “wins”.

For example, consider a piece of silicon that has a donor concentration $N_D = 5\text{x}10^{17}$ cm^{-3} and an acceptor concentration of $N_A = 2\text{x}10^{17}$ cm^{-3}. The empty bonding spots of the acceptor atoms will take up $2\text{x}10^{17}$ cm^{-3} of the electrons from the donors. (The donors compensate the acceptors.) This leaves $3\text{x}10^{17}$ cm^{-3} electrons from the donors to be free electrons, available to carry current. The semiconductor will be n-type. In general, for n-type compensated material the electron concentration is n = N_D - N_A.

Two types of current

Drift current – use an electric field (by applying a voltage). The electric field will move electrons and holes, which is current. This is how current flows in resistors and also in MOSFETs (which are essentially complicated little resistors.

Diffusion current – create a concentration gradient (more electron – or holes – in one place than in another). Through the natural process of the random motion, there will be a net flow of particles from the region of high concentration to the region of lower concentration. This is mechanism for current flow in diodes and bipolar junction transistors.

Both types of current can occur simultaneously. (In fact, drift and diffusion can be handled as two aspects of single driving force, but that sort of approach is beyond us for now.)