Two types of current

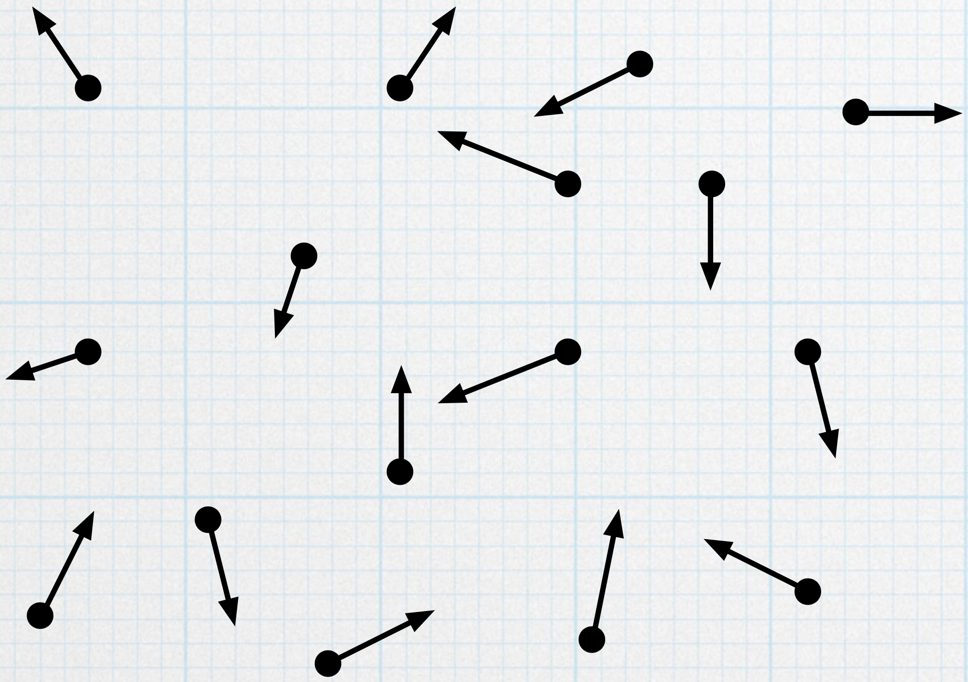
$$J_n = -qnv_n \qquad J_p = qpv_p$$

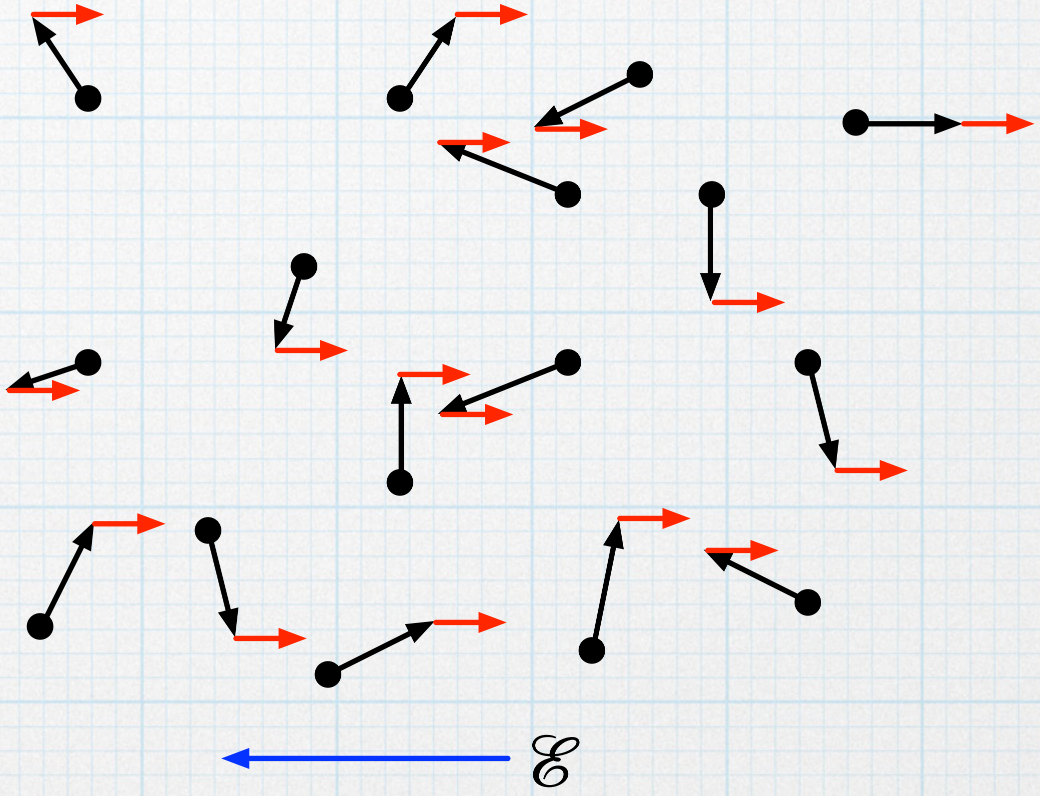
Drift current – An electric field (the result of applying a voltage between two points) is used as a force to push electrons and holes through the semiconductor. This is how current flows in resistors and in field-effect transistors (FETs, which are essentially fancy little resistors). Recall from physics, $\mathscr{E} = -\partial V / \partial x$. Drift current is probably familiar.

Diffusion current – Whenever there is a non-uniform distribution of electrons or holes (a concentration gradient). The process of random motion will cause a net flow of particles from higher concentrations to lower concentrations. This is the mechanism for current flow in diodes and bipolar junction transistors. This is probably a new concept for most EEs/CprEs, but diffusion processes are surprisingly common in the world around us.

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

Consider the free electrons in a *n*-type semiconductor. (All the following arguments apply to holes in *p*-type material, as well.) All electrons are jiggling about, moving a random distribution of directions and speeds. The electrons (and holes) in the semiconductor are very much like atoms or molecules in a gas.



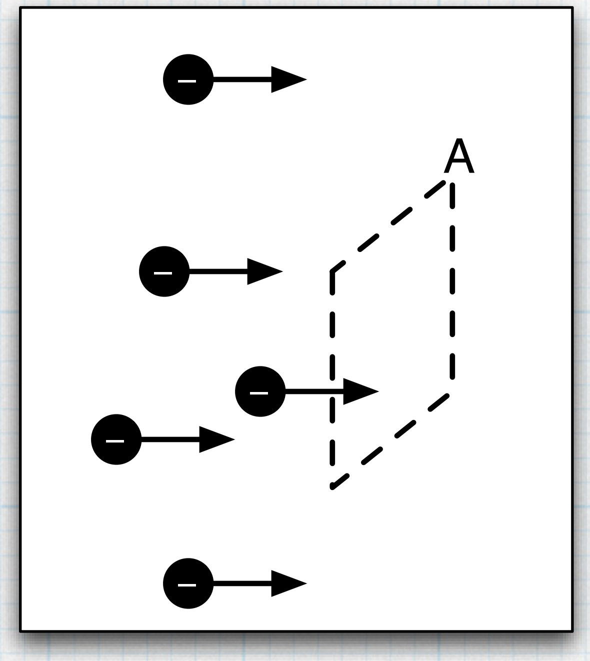


If we look at the entire system, because of the randomness, the *net current* due to the motion is 0. Everything balances out.

If we apply an electric field, the electrons still have random bounces, but each electron gets a nudge to the right. The net electron motion is to the right, which represent a current going to the left.

Flux

The movement of particles can be described using a quantity called flux. Flux is the number crossing through a given area in a given time. Typical units would be $m^{-2}s^{-1}$.



If the particles have a concentration n and are all moving with net velocity v, then within a time interval Δt, the absolute number of particles that moved through the area would be $\Delta N = nvA\Delta t$. Then the flux is

$$\mathscr{F} = \frac{\Delta N}{A\Delta t} = nv \qquad \left[\frac{1}{\mathrm{m}^3}\right]\left[\frac{\mathrm{m}}{\mathrm{s}}\right] = \left[\frac{1}{\mathrm{m}^2\mathrm{s}}\right]$$

If the moving particles are charged (electrons or holes), then the flux represents a current density:

$$q\mathscr{F} = J = qnv \qquad [\mathrm{C}]\left[\frac{1}{\mathrm{m}^2\mathrm{s}}\right] = \left[\frac{\mathrm{C}}{\mathrm{m}^2\mathrm{s}}\right] = \left[\frac{\mathrm{A}}{\mathrm{m}^2}\right]$$

Electrons (holes) in an electric field

Consider a free electron. (The story is the same with holes.) We know that an electric field represents a force applied to a charged particle,

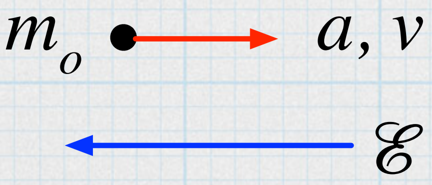
$$F = q\mathcal{E}$$

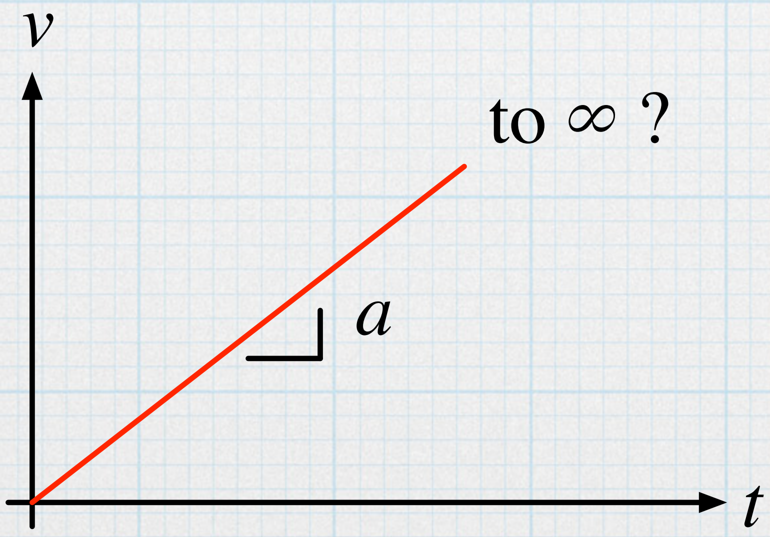
Using one of Newton's Laws, the electric field will cause the charged particle to *accelerate*.

$$F = q\mathcal{E} = m_o a = m_o \frac{dv}{dt}$$

If the electric field is constant in time, the particle will have a constant acceleration, meaning that it speeds up linearly with time.

$$v = \left(\frac{q\mathcal{E}}{m_o}\right) t$$



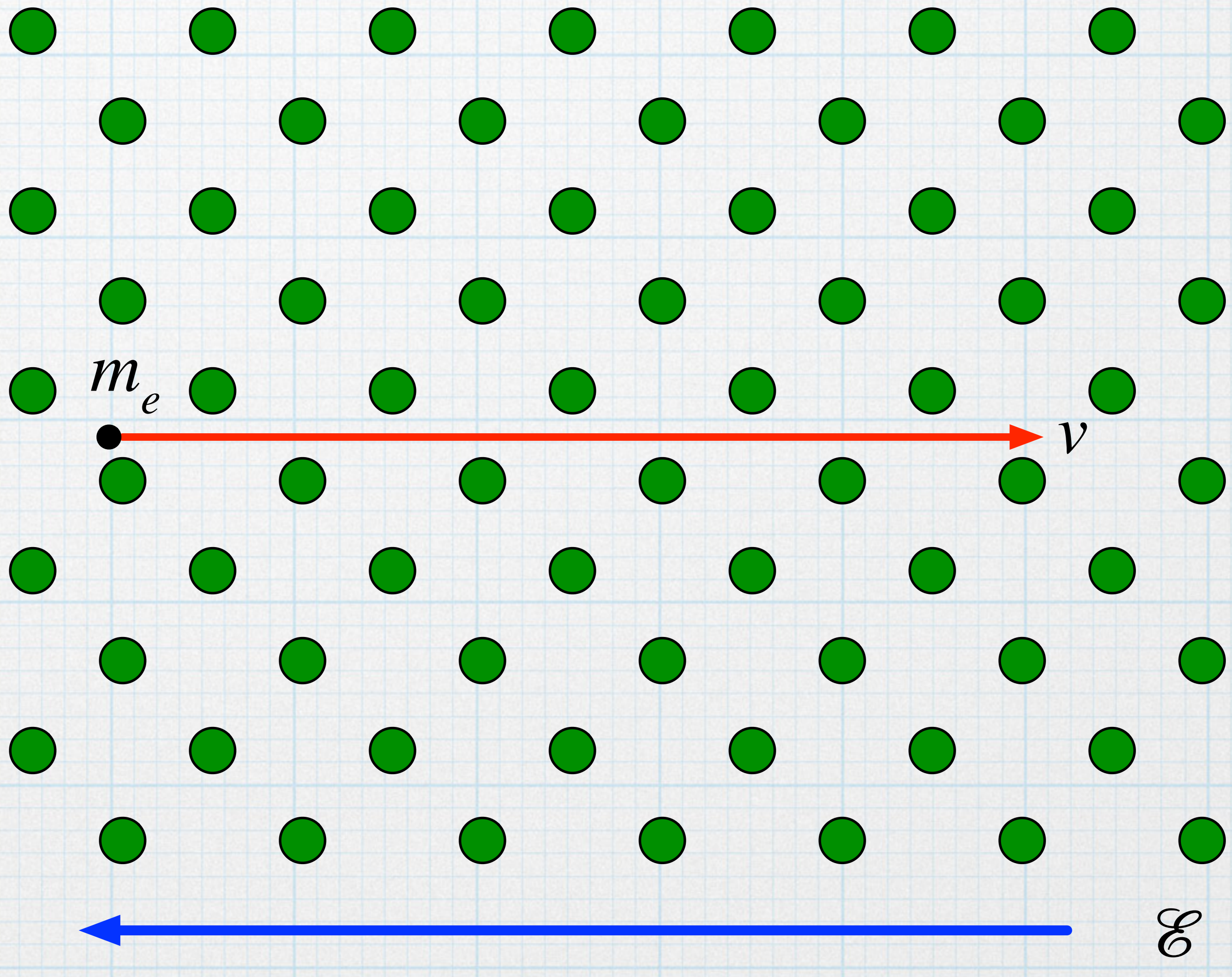


If we have enough energy and time, the electron would eventually hit the speed of light, c_o.

Electron (or hole) in a crystal

For a free electron inside a *perfect* semiconductor crystal with a uniform electric field applied, the situation is a bit surprising (and un-intuitive) — it is exactly the same as the free electron case. A constant electric field causes the electron to accelerate uniformly.

$$v = \left(\frac{q\mathscr{E}}{m_e} \right) t$$

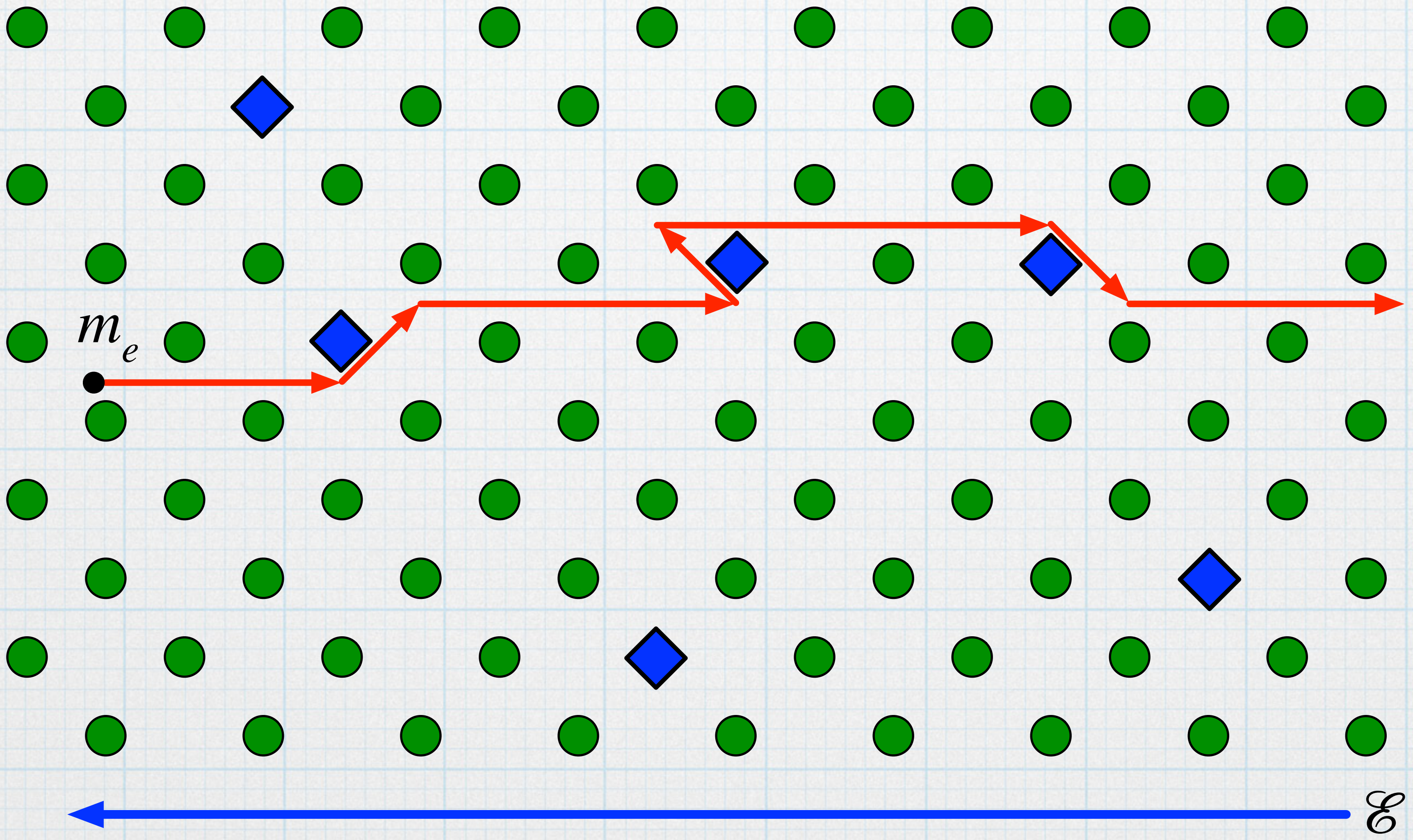


However, the electron seems to move as if it has different mass. We call it the effective mass. (This is very weird, and we are not going to explain it here.) Otherwise, it behave as if it were a free electron.

$$m_e = m^* m_o$$

However, there *are no perfect crystals*. With 10^{22} atoms per cubic centimeter, there will always be some imperfections. Even if the crystal is 99.999999% perfect, (1 flaw per billion atoms), that means the defect density is 10^{13} per cubic centimeter, which is still a rather high number.

Whenever an electron (or hole) encounters an imperfection, it will be deflected off its trajectory.



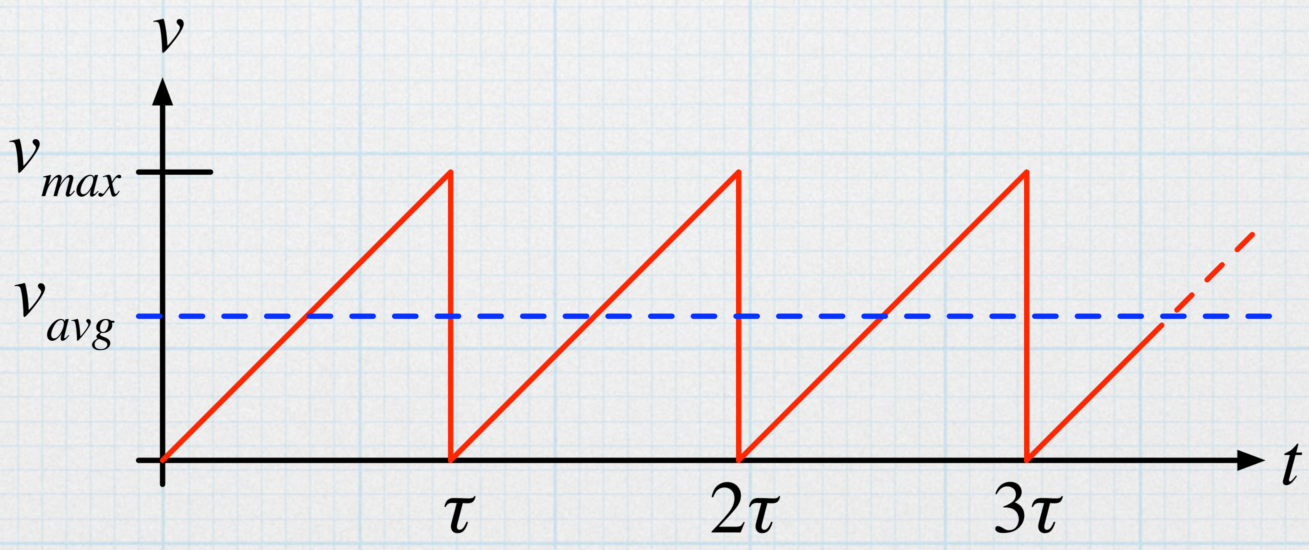
How does scattering affect the current? To get a feel for the situation, we can contrive an overly simplified model of what is happening.

1. A uniform electric field accelerates an electron from rest.
2. After exactly τ seconds, the electron hits an impurity that brings it to a dead stop.
3. Since the $\mathscr{E}$-field is still present, the electron begins to accelerate again.
4. After another τ seconds, the electron hits another impurity and comes to a dead stop again.
5. Keep repeating.

Then find the average value of velocity.

$$v_{max} = \frac{q\mathscr{E}}{m_e}\tau$$

$$v_{avg} = \frac{q\mathscr{E}\tau}{2m_e}$$



This herky-jerky motion is the result of the field accelerating the electrons and the collisions that stop them. The combination leads to an interesting result when looking at the average effect — the *average velocity* is proportional to the field.

$$v_{avg} = \left(\frac{q\tau}{2m_e} \right) \mathscr{E}$$

The proportionality constant is a property of the sample — the time between collisions and the effective mass of the electrons. We call this quantity the *mobility*. It tells how fast the electrons will *drift* when an electric field is applied.

$$v_{avg} = \mu \mathscr{E} \qquad \left[\frac{\mathrm{m}}{\mathrm{s}} \right] = \mu \left[\frac{\mathrm{V}}{\mathrm{m}} \right] \qquad \mu \rightarrow \left[\frac{\mathrm{m}^2}{\mathrm{V} \cdot \mathrm{s}} \right] \text{ or } \left[\frac{\mathrm{cm}^2}{\mathrm{V} \cdot \mathrm{s}} \right]$$

Since mobility depends on effective mass, it is different for electrons and holes in a given semiconductor. Also, effective masses differ by semiconductor — some are high some are low. And mobility depends on how long the electron or hole can travel between collisions, τ.

Scattering & mobility

Any imperfection in the periodicity of the crystal will cause carriers to scatter.

- charged impurity atoms substituted into crystal locations
- vibrations in the crystal
- other carriers (a free electron or a hole is a type of imperfection)
- dislocations (lattice planes shifted out of place)
- surfaces (can be viewed as a severe form of dislocation)
- uncharged impurity atoms substituted into the crystal
- vacancies or interstitial atoms (missing or misplaced atoms)

Any given sample of semiconductor crystal will probably have all of these scattering mechanisms present and all will contribute to the scattering that limits the speed of the carriers. The relative contribution of each depends on the relative amounts. The more scatterers, the shorter the time between collisions and the lower the mobility. The first two items in the list are usually the most important in a semiconductor.

Each type of scattering mechanism will have its own average time between collisions. The effect with the smallest τ will have the biggest effect in limiting mobility.

Analogy

When walking down a city sidewalk, you might have to walk around – scatter off – a number of things: other pedestrians, lamp posts, mailboxes, discarded scooters, open manholes, etc. The more you have to weave and dodge, the slower your average velocity down the sidewalk. During the day, when the streets are crowded, avoiding other pedestrians is the biggest limit in determining how fast you can walk. Avoiding the other items is a relative minor annoyance. The τ of "colliding" with other people is much smaller than the τ for the other obstacles. But at night, when the streets are empty of people, the τ for avoiding pedestrians is much bigger than during the day – there aren't as many to avoid. In that case, the mean free times for the lamp posts, mailboxes, etc. might become the limits on the speed at which you can travel down the sidewalk.

So it is with electrons and holes – every imperfection will cause them to scatter, but some are more important than others, due mainly to the effective concentrations. In a typical semiconductor, charged impurity atoms and phonons are usually the most prevalent types of scattering, and so those two are most important in determining the overall mobility.

To combine the effects of different scattering mechanisms, we need a composite scattering time that includes everything. Because shorter times are more limiting and thus more important, we should "add" the times using inverses (known as Matthiessen's rule):

$$\frac{1}{\tau_{total}} = \frac{1}{\tau_1} + \frac{1}{\tau_2} + \frac{1}{\tau_3} + \ldots$$ (Looks like parallel resistors.)

The most frequency scattering (i.e. having the shortest scattering time) will have the biggest effect in determining the overall mobility. Using the individual scattering times, we can define a separate mobility for each scattering mechanism, $\mu_1 = q\tau_1/m$, $\mu_2 = q\tau_2/m$, etc. and we can re-cast above equation using mobilities.

$$\frac{1}{\mu_{total}} = \frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{1}{\mu_3} + \ldots$$

Example: For electrons in some semiconductor, two scattering mechanisms lead to two individual mobilities $\mu_1 = 500\ \mathrm{cm^2/V{\cdot}s}$ and $\mu_2 = 1000\ \mathrm{cm^2/V{\cdot}s}$. What is the overall mobility for electrons in the semiconductor?

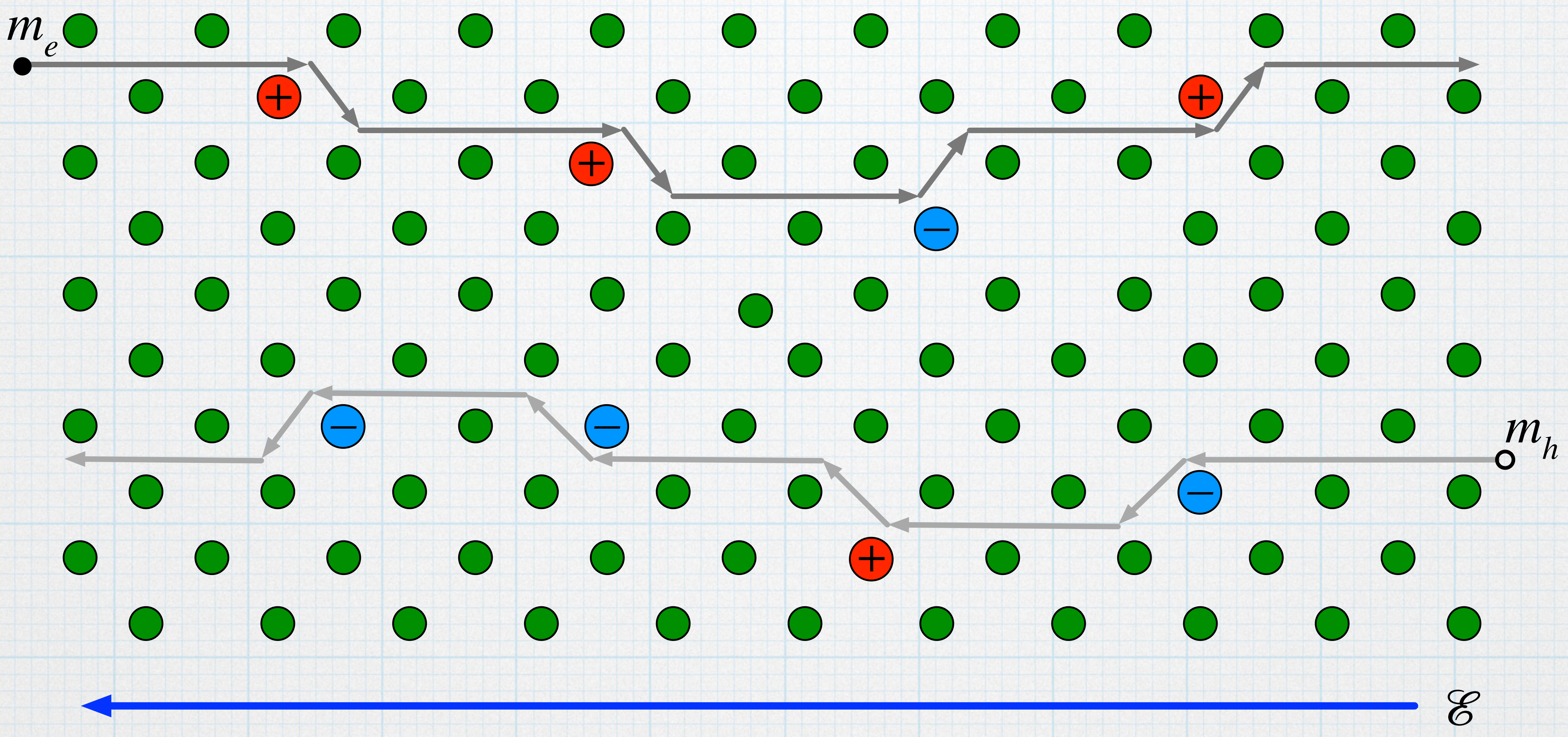
$$\mu_{total} = \left[\frac{1}{\mu_1} + \frac{1}{\mu_2}\right]^{-1} = \left[\frac{1}{500\ \frac{\mathrm{cm^2}}{\mathrm{V \cdot s}}} + \frac{1}{1000\ \frac{\mathrm{cm^2}}{\mathrm{V \cdot s}}}\right]^{-1} = 333\ \frac{\mathrm{cm^2}}{\mathrm{V \cdot s}}$$

Ionized impurity scattering

One of the two most important sources of scattering in a semiconductor is ionized impurity scattering. If a lattice atom becomes ionized, it will cause passing electrons or holes to deflect in some fashion (like charges repel, opposite attract.) What is the source of the ionized lattice atoms? They are introduced intentionally! They are an unavoidable consequence of the doping process. When a phosphorus donor atom substitutes for a silicon atom, its fifth electron breaks loose and moves away. The atom is short one electron to match the positive charge in the nucleus, and so the atom ends up with a $+q$ net charge. This charge is immobile – it is locked into the lattice. The story is the same with a boron acceptor atom, except that the boron accepts an electron from somewhere else in the crystal, resulting in a net charge of $-q$, also locked in place.

Any electron or hole passing near an ionized phosphorus or boron atom in the crystal will be deflected (scattered). Increasing the doping to increase the carrier concentration causes the mobility will go down. There are more carriers, but they are moving slower. This seems a bit ironic — and perhaps self-defeating since it appears that the two effects are nullifying each other. However, we will see shortly that while ion impurity scattering certainly has an effect, it doesn't completely cancel the changes in carrier concentration.

Ionized impurity scattering



Electrons and holes are scattered by both positive and negative ions.

Scattering from lattice vibrations (phonons)

The other major semiconductor scattering mechanism is caused by lattice vibrations. In simple terms, the atoms in the crystal are not static but are constantly oscillating around the nominal lattice position. This is an acoustic (mechanical) vibration akin to a movement of a plucked guitar string, but on a much smaller scale. From quantum mechanics, any oscillating wave, including acoustic waves, can be treated as a particle. In the case of acoustic waves, the corresponding particles are known a phonons.

Electrons and holes can be scattered very effectively from the lattice vibrations (which could also be called colliding with a phonon). There will be a scattering time, τ_{phonon}, and a corresponding mobility, $\mu_{phonon} = q\tau_{phonon}/m_e$.

Lattice vibrations are always present in a semiconductor (or in any solid material). The atoms will always be jiggling due to the thermal energy of the material. In fact, lattice vibrations / phonons are a way to describe the thermal energy. When a solid material heats up, the atoms vibrate more vigorously and when the material cools down, they vibrate less. Viewed the opposite way, the stronger vibrations of the atoms, the hotter the material must be.

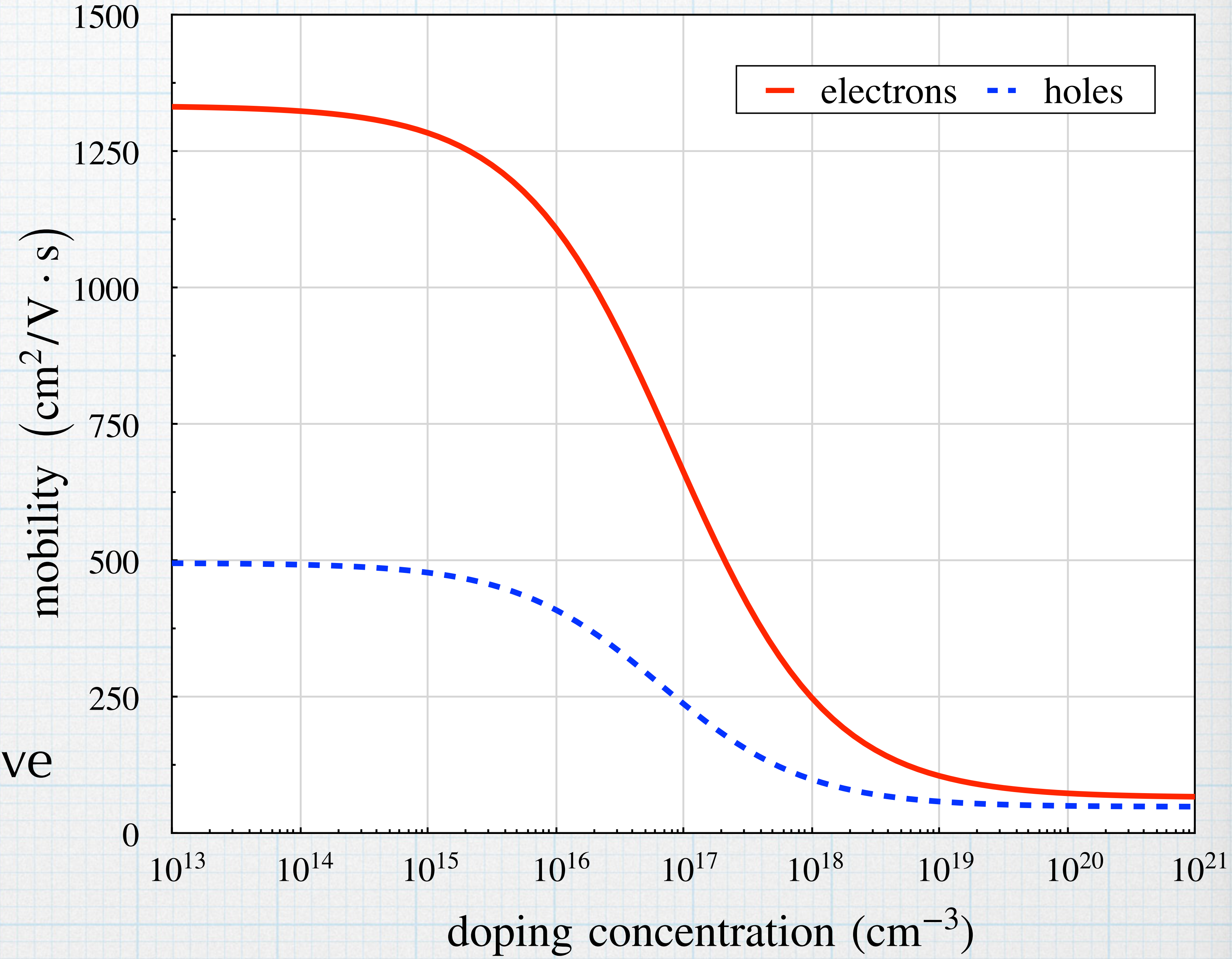
Obviously, scattering from phonons is very temperature-dependent. Cooling the semiconductor will speed up the carriers! On the other hand, phonon scattering is essentially independent of doping level.

Mobility as a function of doping concentration

$$\mu = \left[\frac{1}{\mu_{ion}(N)} + \frac{1}{\mu_{phonon}} \right]^{-1}$$

Phonons dominate at low doping, ionized impurities become dominate as doping increases.

Equations below are curve fits to experimental data for electrons and holes.



$$\mu_n = 65\,\frac{\text{cm}^2}{\text{V}\cdot\text{s}} + \frac{1268\,\frac{\text{cm}^2}{\text{V}\cdot\text{s}}}{1 + \left(\frac{N_D}{8.5\times10^{16}\,\text{cm}^{-3}}\right)^{0.71}}$$

$$\mu_p = 48\,\frac{\text{cm}^2}{\text{V}\cdot\text{s}} + \frac{447\,\frac{\text{cm}^2}{\text{V}\cdot\text{s}}}{1 + \left(\frac{N_A}{6.63\times10^{16}\,\text{cm}^{-3}}\right)^{0.76}}$$

Drift current and conductivity

$$J_n = -\,qnv_n = -\,qN_Dv_n \qquad J_p = qpv_p = qN_Av_p$$

$$v_n = -\,\mu_n\mathscr{E} \qquad v_p = \mu_p\mathscr{E}$$

$$J_n = qn\mu_n\mathscr{E} = \left(q\mu_nN_D\right)\mathscr{E} \qquad J_p = qn\mu_p\mathscr{E} = \left(q\mu_pN_A\right)\mathscr{E}$$

$$J_n = \sigma_n\mathscr{E} \qquad J_p = \sigma_p\mathscr{E}$$

$\sigma_n = q\mu_n n$ (or $q\mu_p p$)is the *conductivity* of the material – it is a property of a given semiconductor with a given doping concentration.
(Observation: $J = \sigma\mathscr{E}$ is the microscopic version of Ohm's Law!)

Looking at units:

$$\left[\frac{\mathrm{A}}{\mathrm{m}^2}\right] = \sigma\left[\frac{\mathrm{V}}{\mathrm{m}}\right] \qquad \sigma \rightarrow \left[\frac{A}{\mathrm{V}\cdot\mathrm{m}}\right] = \left[\frac{1}{\Omega\cdot\mathrm{m}}\right] = \left[\frac{\mathrm{S}}{\mathrm{m}}\right] \quad \text{or} \quad \left[\frac{\mathrm{S}}{\mathrm{cm}}\right]$$

Typical value: $\sigma_n = \left(1.6\times10^{-19}\,\mathrm{C}\right)\left(1000\,\frac{\mathrm{cm}^2}{\mathrm{V}\cdot\mathrm{s}}\right)\left(10^{17}\,\mathrm{cm}^{-3}\right) = 16\,\frac{\mathrm{S}}{\mathrm{cm}}$

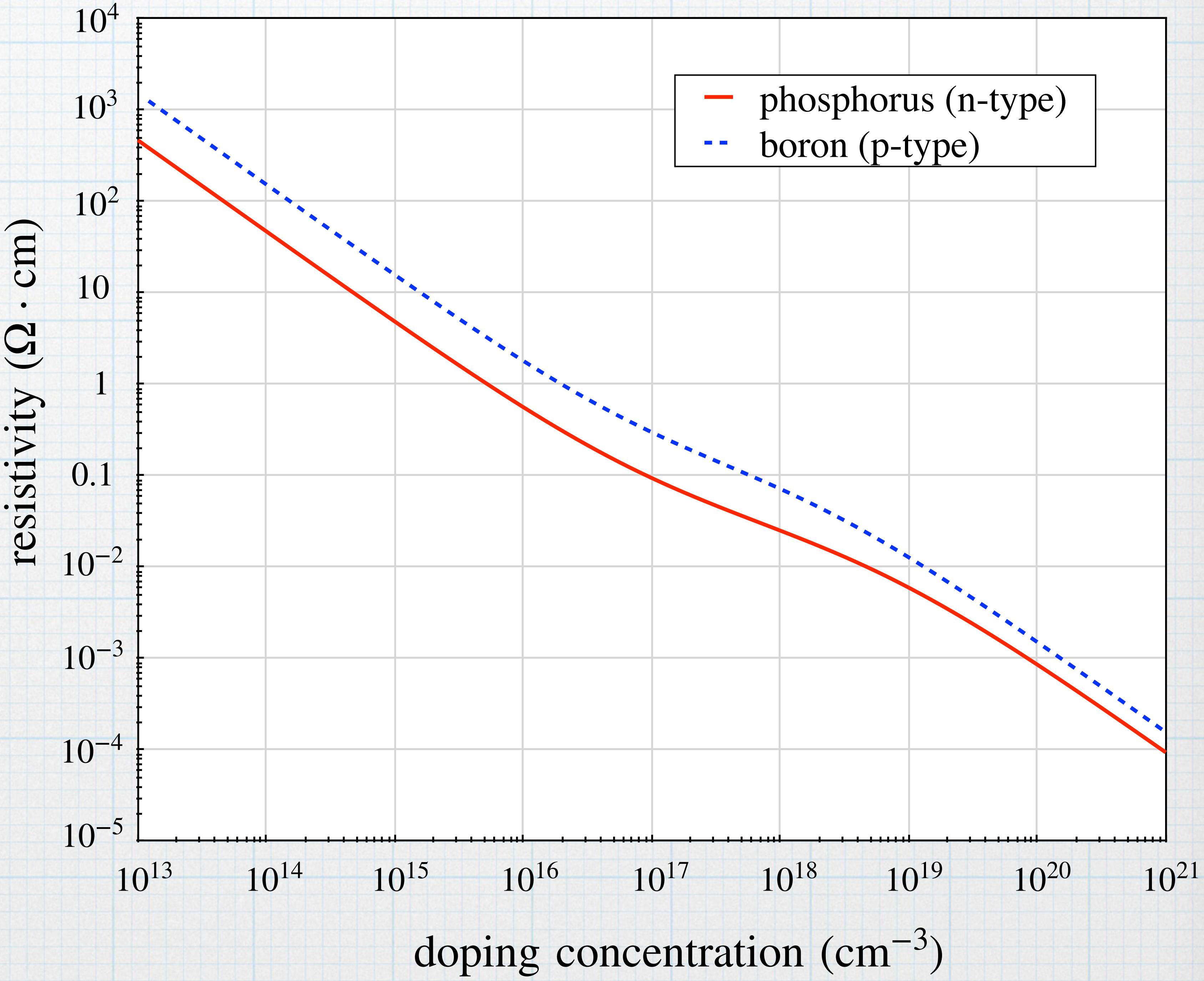
Resistivity

A related quantity is resistivity: $\rho_n = 1/\sigma_n = (q\mu_n n)^{-1}$ [$\rho_p = (q\mu_p p)^{-1}$ for p-type material]. No new information here, but often resistivity will be reported rather than conductivity.

$$J_n = \frac{\mathcal{E}}{\rho_n} \qquad J_p = \frac{\mathcal{E}}{\rho_p}$$

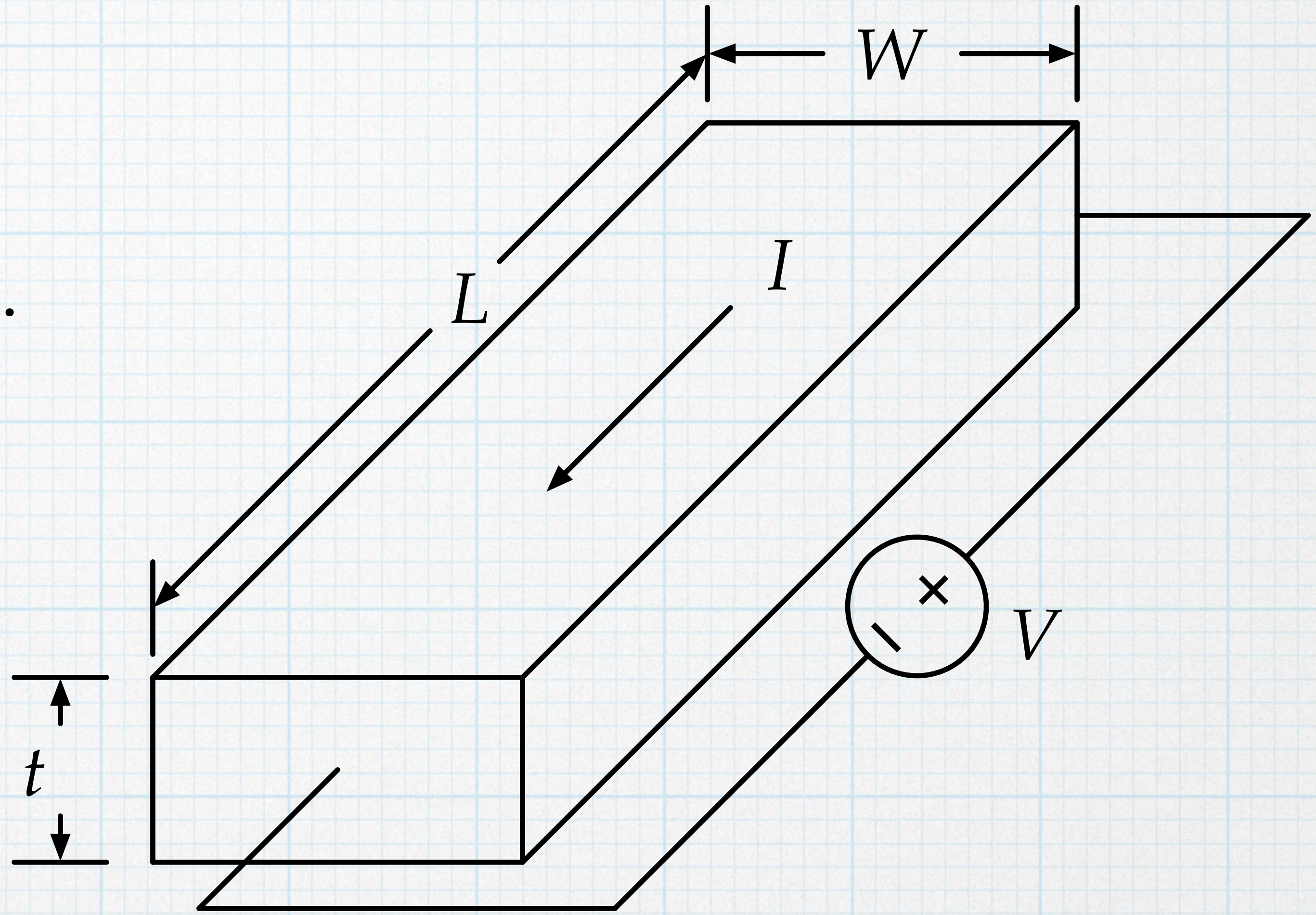
It is common to plot resistivity vs. doping concentration, using the resistivity definition and the mobility from page 15.

Of course, for the same concentrations, holes have different mobilities and so will have different resistivities. Also, we can see the slight wiggles in the curves due to the changing mobility.



Resistivity → resistors

Make a rectangular bar resistor.
Apply a voltage down the length.



$$J_n = \frac{\mathscr{E}}{\rho_n}$$

$$\frac{I}{W \cdot t} = \frac{1}{\rho_n}\frac{V}{L}$$

$$I = \frac{Wt}{\rho_n L}V$$

$$I = \frac{V}{R}$$

$$R = \frac{\rho L}{Wt} = \frac{L}{q\mu_n N_D W t}$$

Classic resistor equation. The properties of the "device" are a combination of material properties (μ_n and N_D) and geometrical shape (L, W, t).

Example

A piece of silicon is doped n-type with phosphorus with a doping level of N_D = 3x10^{17} cm^{-3} (3x10^{23} m^{-3}). The corresponding electron mobility is μ_n = 750 cm^2/V·s (0.0750 m^2/V·s) What is the conductivity of the n-type sample?

The electron concentration is equal to the doping concentration.

$$n = N_D = 3\text{x}10^{17}\ \text{cm}^{-3}.$$

Then: $\sigma = q\mu_n n = (1.6\text{x}10^{-19}\ \text{C})(750\ \text{cm}^2/\text{V}\cdot\text{s})(3\text{x}10^{17}\ \text{cm}^{-3}) = 36\ \text{A}/\text{V}\cdot\text{cm}.$

$$= 36\ \text{S}/\text{cm} = 36\ \Omega^{-1}\text{cm}^{-1}$$

What is the resistivity?

$$\rho = \frac{1}{\sigma} = \frac{1}{36\ \text{S}/\text{cm}} = 0.028\ \Omega \cdot \text{cm}$$

Example 2

A piece of silicon is doped *p*-type with born with a doping level of N_A = 2x10^{16} cm^{-3} (2x10^{22} m^{-3}). The corresponding hole mobility is μ_p = 367 cm^2/V·s (0.0367 m^2/V·s) What is the resistivity of the *p*-type sample?

The hole concentration is equal to the doping concentration.

$$p = N_A = 2\text{x}10^{16}\ \text{cm}^{-3}.$$

Then:

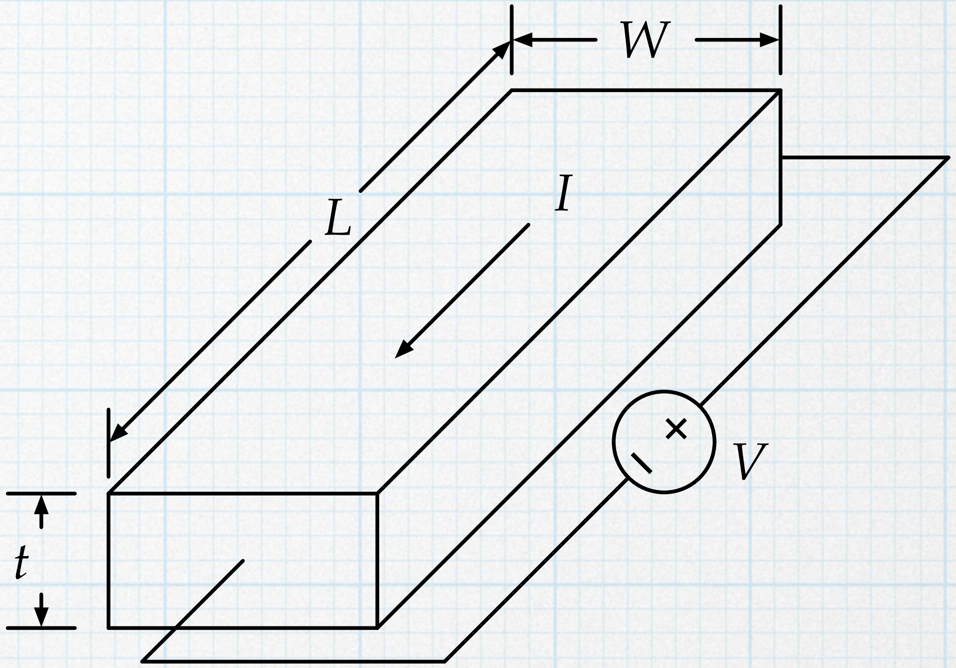
$$\rho = \frac{1}{q\mu_p N_A} = \frac{1}{\left(1.6\times 10^{-19}\,\text{C}\right)\left(367\frac{\text{cm}^2}{V\cdot s}\right)\left(2\times 10^{16}\,\text{cm}^{-3}\right)} = 0.85\,\Omega\cdot\text{cm}$$

What is the conductivity?

$$\sigma = \frac{1}{\rho} = 1.17\ \text{S/cm}$$

Example 3

A resistor is made out of *n*-doped silicon. The doping level is $N_D = 1\text{x}10^{18}\ \text{cm}^{-3}$. (The corresponding electron mobility is $\mu_p = 425\ \text{cm}^2/\text{V}\cdot\text{s}$. The resistor has dimensions of $L = 100\ \mu\text{m}$, $W = 20\ \mu\text{m}$, and $t = 2\ \mu\text{m}$. What is the resistance of the resistor?



Note: 1 μm = 10^{-4} cm = 10^{-6} m.

First, find the resistivity of the *n*-type silicon. $n = N_D = 1\text{x}10^{18}\ \text{cm}^{-3}$.

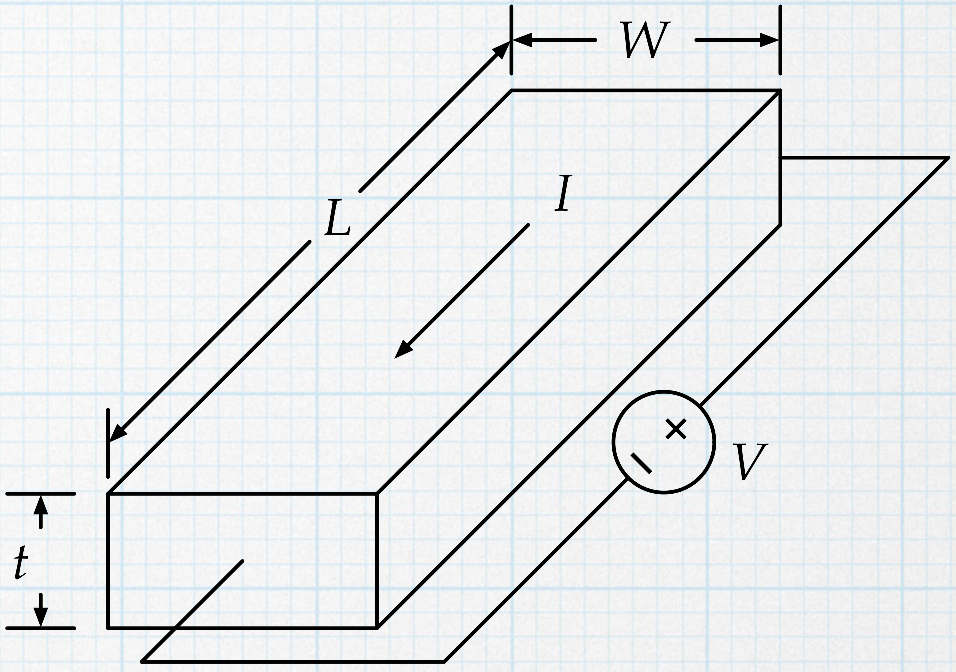
$$\rho = \frac{1}{q\mu_n n} = \frac{1}{(1.6 \times 10^{-19}\ \text{C})\left(425\frac{\text{cm}^2}{V\cdot s}\right)(10^{18}\ \text{cm}^{-3})} = 0.015\ \Omega\cdot\text{cm}$$

Then

$$R = \frac{\rho L}{W \cdot t} = \frac{(0.015\ \Omega\cdot\text{cm})(0.01\ \text{cm})}{(0.002\ \text{cm})(0.0002\ \text{cm})} = 375\Omega$$

Example 4

A resistor is made out of *p*-doped silicon. The doping level is $N_A = 5\text{x}10^{16}\ \text{cm}^{-3}$. The corresponding hole mobility is $\mu_p = 410\ \text{cm}^2/\text{V}\cdot\text{s}$. The cross-sectional area of the resistor is W·t = (5 μm)(1μm) = 5 μm² = $5\text{x}10^{-8}\ \text{cm}^2$. What length is required for a resistance of 5 kΩ?



First, find the resistivity of the *p*-type silicon. $p = N_A = 5\text{x}10^{16}\ \text{cm}^{-3}$.

$$\rho = \frac{1}{q\mu_p p} = \frac{1}{(1.6 \times 10^{-19}\ \text{C})\left(410\frac{\text{cm}^2}{V\cdot s}\right)(5 \times 10^{16}\ \text{cm}^{-3})} = 0.30\ \Omega \cdot \text{cm}$$

$$R = \frac{\rho L}{W \cdot t}$$

$$L = \frac{R \cdot W \cdot t}{\rho} = 8.33 \times 10^{-4}\ \text{cm} = 8.33\mu\text{m}$$ (Pay attention to units.)

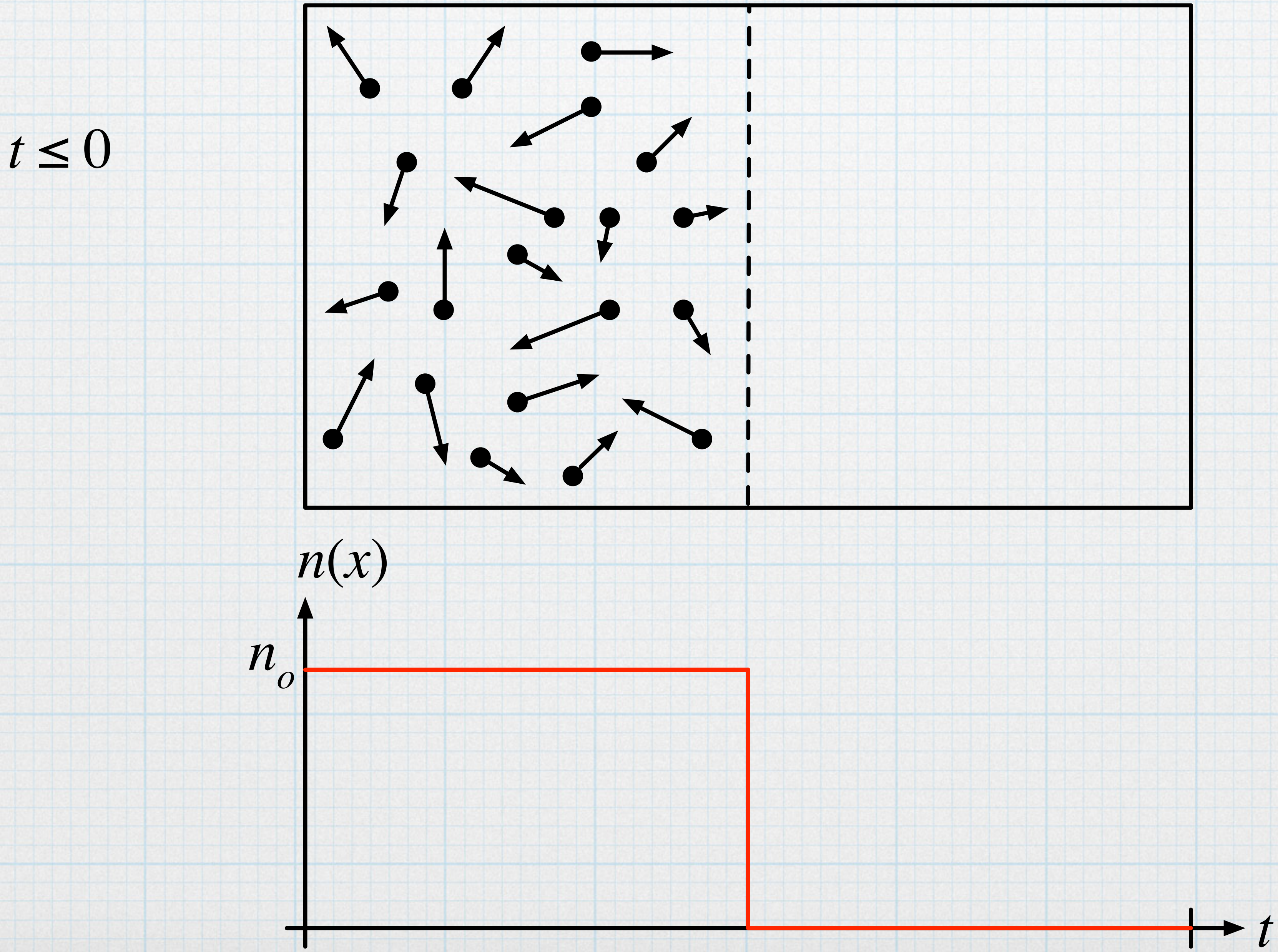
Diffusion

Diffusion is the second method by which electrons and holes can have net current flow. Diffusion is the “smoothing” process by which randomly moving particles will be re-distributed into a more uniform pattern. Whenever particles have a non-uniform distribution and they can move randomly, there will be a net flow from the regions of higher concentration into regions of lower concentration. Diffusion is common process that we can see all the time in the world around us. Here are a few examples:

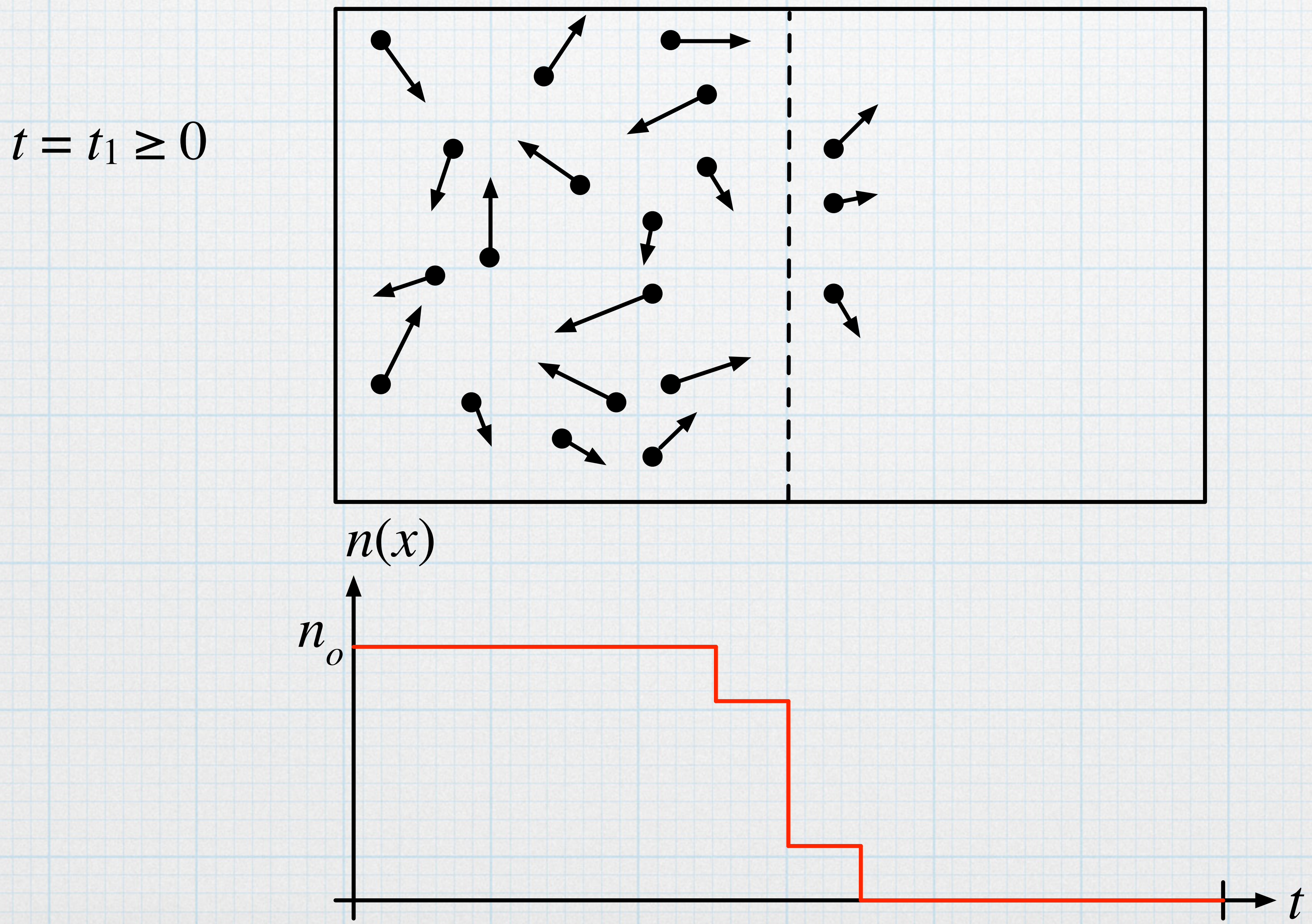
- steam billowing off a pot of boiling water
- oders moving through the air to be detected by your nose
- tea moving from tea bag into the surrounding water
- an ink stain moving through a piece of cloth
- heat moving through a solid.

The random motion of electrons and holes in a semiconductor leads to diffusion currents.

To develop some qualitative understanding of diffusion, we can contrive a starting concentration gradient of some particles and see what happens as they move around. Consider a box containing particles These could be atoms/molecules in the air or electrons/holes in a semiconductor. Initially, the particles are confined to the left side of the box, but are evenly distributed on that side. The right side is empty. Clearly, there is a strong concentration gradient.

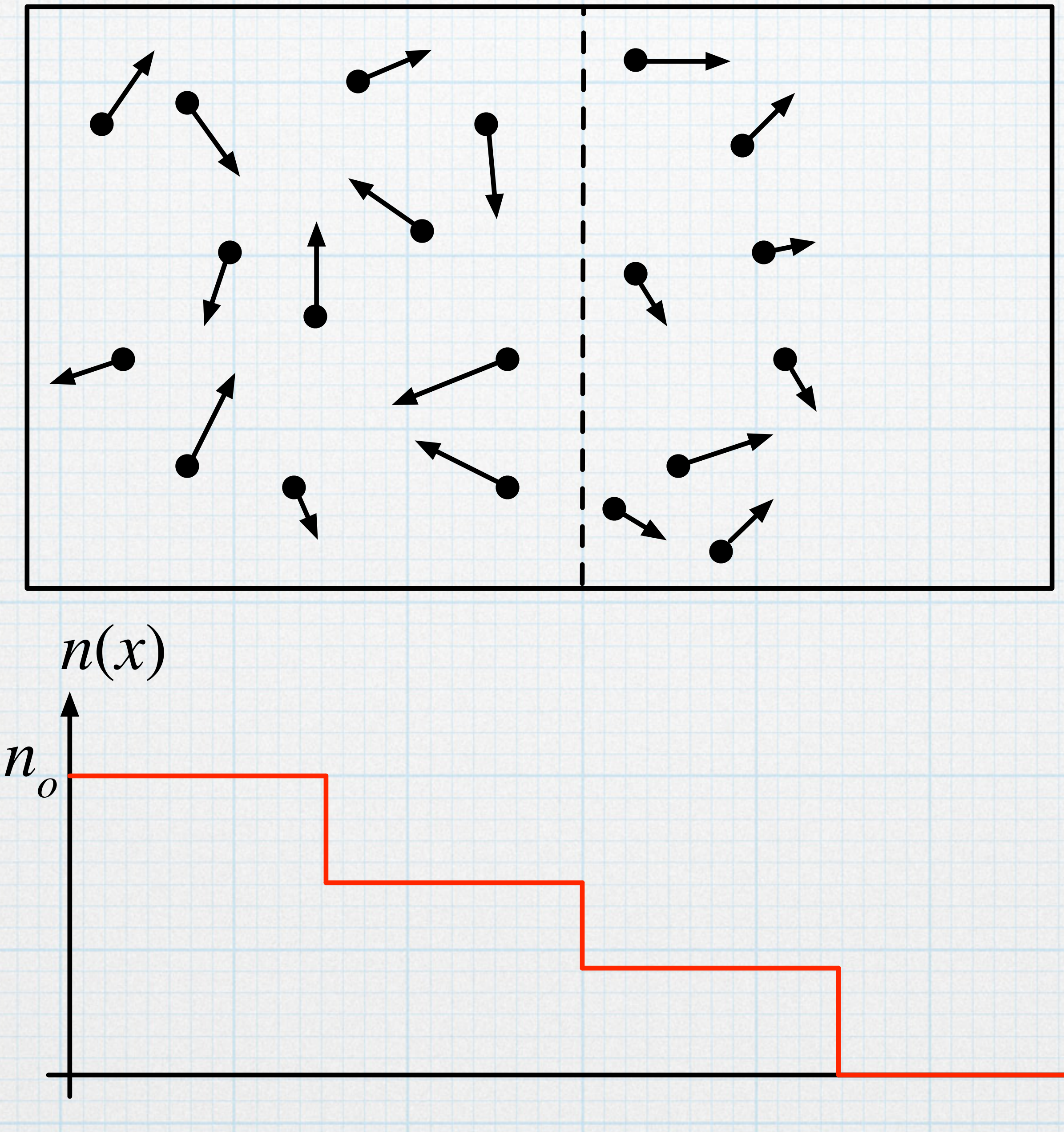


At $t = 0$, the barrier confining the particles to the left side is removed. The particles are still bouncing around in the same random manner, but the particles that were near the barrier and moving towards the right can now move into the empty space. However, there are no particles available to move to the left to balance those out. There is a net flow of particles to the right across the former dividing line.

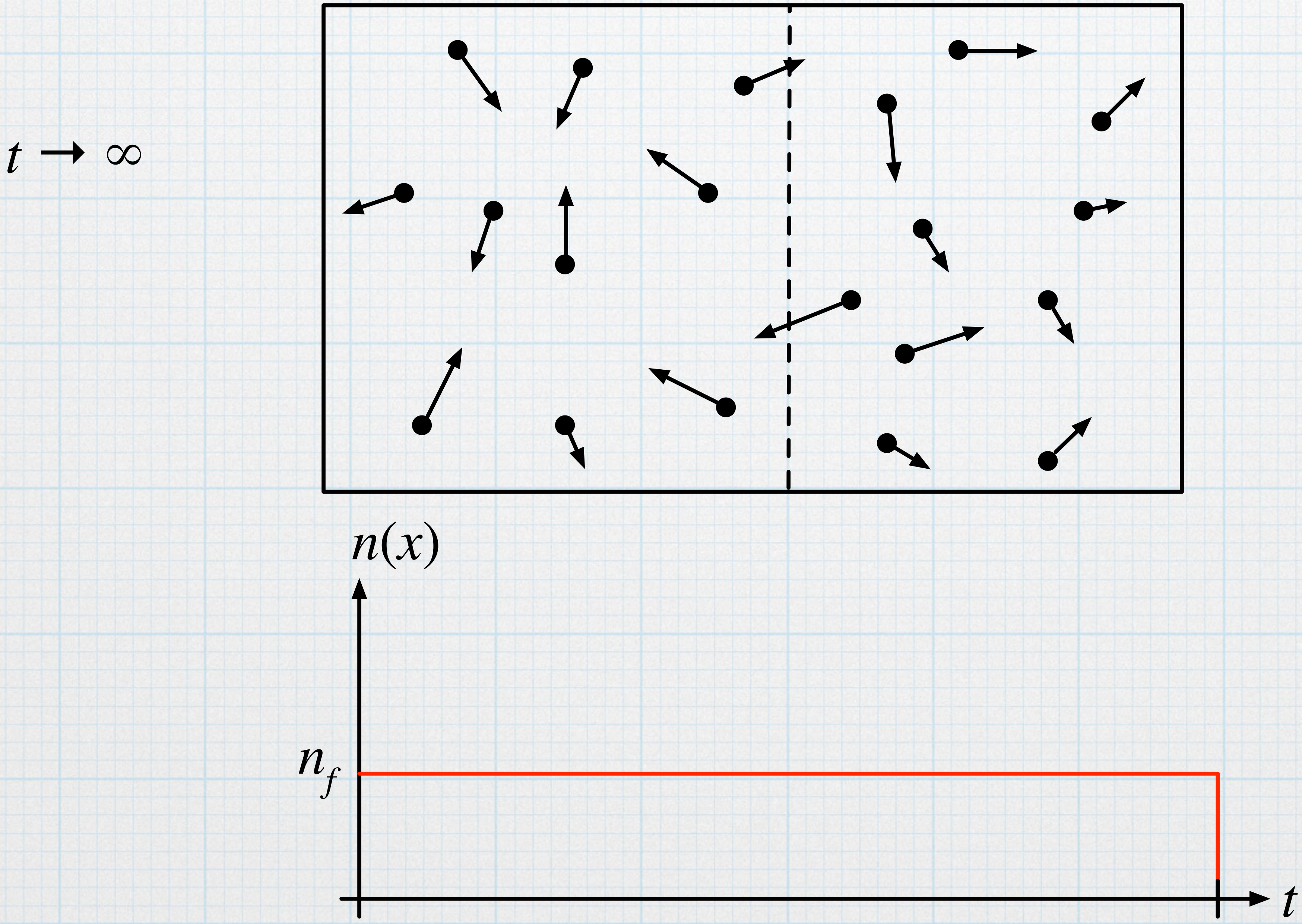


As time increases, the net flow from left to right continues – as long as there are more particles on the left than there are on the right.

$t = t_2 \geq t_1$



Eventually, the particles are re-distributed back to a uniform, but lower, concentration throughout the entire box. Since the distribution is uniform and the motion of the individual particles is still random, there is no longer a net flow in any direction – everything averages to zero again.



Fick's Law

Adolf Fick came up with a quantitative description of the diffusion process (1855). If the process is described in words, the equation is easy to understand: A gradient in the concentration of particles leads to a proportional net flow (i.e. a flux) of particles from the higher concentration to the lower concentration:

$$\mathscr{F} = -D\frac{\partial N}{\partial x}$$

where D is the diffusion coefficient – a macroscopic quantity that describes microscopic movement of the individual particles. It depends on how easily the particles can move and how much they bounce around (scatter) while they are moving. If $D = 0$, there can be no diffusion. (Consider the water molecules in a block of ice versus the molecules in a cup of liquid.) The diffusion coefficient is a property of the particles in whatever medium they are moving. The negative sign accounts for the movement from high to low concentration.

$$\left[\frac{1}{\mathrm{m}^2 \cdot \mathrm{s}}\right] = D\left[\frac{\frac{1}{\mathrm{m}^3}}{\mathrm{m}}\right] \qquad D \rightarrow \left[\frac{\mathrm{m}^2}{\mathrm{s}}\right] \quad \text{or} \quad \left[\frac{\mathrm{cm}^2}{\mathrm{s}}\right]$$

Diffusion current density

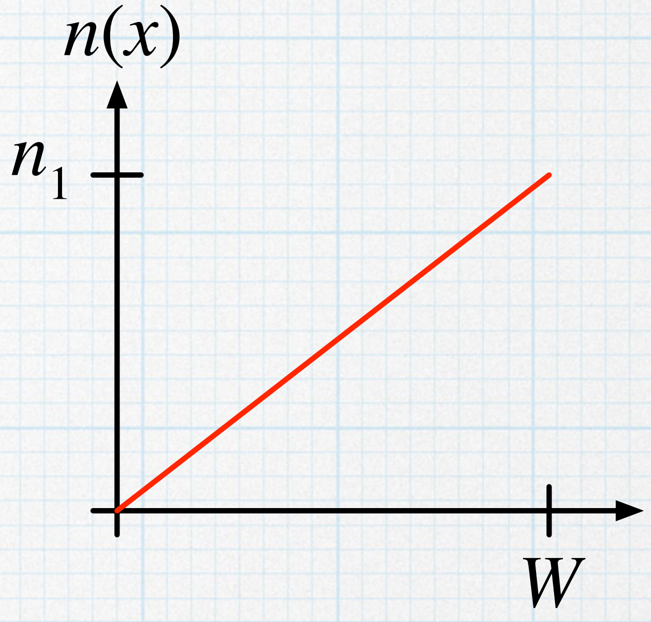
If the particles are charged (like electrons or holes) then the diffusion is an electrical current.

$$J_n = -q\mathscr{F}_n = qD_n \frac{\partial n(x)}{\partial x}$$

$$J_p = q\mathscr{F}_p = -qD_p \frac{\partial p(x)}{\partial x}$$

This gives us another mechanism for making currents flow in a semiconductor – artificially create a concentration gradient and let the electrons or holes diffuse in response. The trick is in creating the concentration gradient. We will see shortly that joining together p and n regions leads to concentration gradients that can be manipulated. This is how diodes and bipolar junction transistors are made.

Example



The electron distribution in some region of a piece of silicon region has a linear variation:

$$n\left(x\right) = n_1 \frac{x}{W}$$

where $n_1 = 10^{12}\ \text{cm}^{-3}$ and $W = 2\ \mu\text{m}$. The diffusion coefficient for the electrons is 35 cm^2/s.

$$\frac{\partial \left(x\right)}{\partial x} = \frac{n_1}{W}$$

$$J_n = qD_n \frac{\partial n\left(x\right)}{\partial x} = qD_n \frac{n_1}{W}$$

$$= \left(1.6 \times 10^{-19}\,\text{C}\right)\left(35\frac{\text{cm}^2}{s}\right)\left(\frac{10^{12}}{2 \times 10^{-4}\,\text{cm}}\right) = 28\ \frac{\text{mA}}{\text{cm}^2}$$

The electrons are moving to the left, and the current is to the right.

Example

The hole distribution in the region x > 0 of a piece of silicon has an exponential variation:

$$p(x) = p_1 \exp\left(-\frac{x}{L}\right)$$

where $p_1 = 10^{14}\ \mathrm{cm}^{-3}$ and $L = 5\ \mu\mathrm{m}$. The diffusion coefficient for the holes is $10\ \mathrm{cm}^2/\mathrm{s}$.

$$\frac{\partial p(x)}{\partial x} = -\frac{p_1}{L} \exp\left(-\frac{x}{L}\right)$$

$$J_p(x) = -qD_p \frac{\partial p(x)}{\partial x}$$

$$= qD_p \frac{p_1}{L} \exp\left(-\frac{x}{L}\right)$$

$$= \left(1.6 \times 10^{-19}\,\mathrm{C}\right)\left(10\frac{\mathrm{cm}^2}{\mathrm{s}}\right)\left(\frac{10^{14}\,\mathrm{cm}^{-3}}{5 \times 10^{-4}\,\mathrm{cm}}\right) \exp\left(-\frac{x}{L}\right)$$

$$= \left(320\,\frac{\mathrm{mA}}{\mathrm{cm}^2}\right) \exp\left(-\frac{x}{L}\right)$$