302L s20 lecture 14 - batteries and resistance

1 Summary

- A battery is a device that exploits a chemical reaction to generate a potential difference between its positive and negative terminals.

- Depletion of charge from the terminals of a battery induces further chemical reaction which acts to restore the terminals to their equilibrium potential.

- If charge is drawn from the battery at a sufficiently slow rate, then the battery is able to maintain its equilibrium potential at its terminals and the battery can be considered a source of constant electric potential (voltage).

- Conceptually, a resistor is a piece of material that conducts electricity poorly. A resistor has two ends where charge can enter or exit.

- A charge carrier in a resistor suffers frequent collisions with its environment. As a result, a charge carrier in a resistor under the influence of a constant electric field $\vec{E}$ travels on average with a *drift velocity* $\vec{v}_d$ given by the product $\tau \frac{q\vec{E}}{m}$ where τ is the average time between collisions, q is the charge of the charge carrier, and m is the mass of the charge carrier.

- The *current* (denoted I) through a resistor is the ratio $\frac{\Delta Q}{\Delta t}$ of the amount of charge ΔQ entering (or exiting) the resistor over a time interval of duration Δt. Conceptually current is a flow of electric *charge*.

- The SI unit of current is the ampere (or amp for short), denoted A, and $1\,\text{A} = 1\,\text{C/s}$.

- To determine the *sign* of the current (i.e. positive or negative) through a resistor, we must first pick one end of the resistor. If a net positive charge is entering this end, we say the current is positive.

- *Ohm's law:* The current I through a resistor is proportional to difference V in electric potential between the two ends of the resistor. The proportionality constant, denoted R, is known as the resistor's *resistance*. In short:
$$V = IR$$

  Conceptually, at a given voltage a resistor with a large resistance will conduct a weaker current than one with a small resistance.

- The SI unit of resistance is the *ohm*, denoted Ω, and $1\,\Omega = 1\,\text{V/A}$

- A resistor with a cross sectional area A and length l has a resistance $R = \rho\frac{l}{A}$. ρ is a proportionality constant associated with the material composing the resistor. It has SI units of $\Omega\,\mathrm{m}$ and is known as the *resistivity* of the material.

- A material containing charge carriers of charge q and mass m with a mean time τ between collisions has a resistivity ρ given by

$$\rho = \frac{m}{nq^2\tau}$$

where n is the number of charge carriers per unit volume (i.e. the charge carrier *density*). Note how the different microscopic properties of the material influence its macroscopic resistivity.

2 The Voltaic Cell (Battery)

In discussing the topic of electric current it will be helpful to have at our disposal a source of constant *voltage*. So far we have discussed the capacitor, which has an electric potential difference V between its two plates given by the ratio Q/C of the charge $\pm Q$ on its plates and its capacitance C:
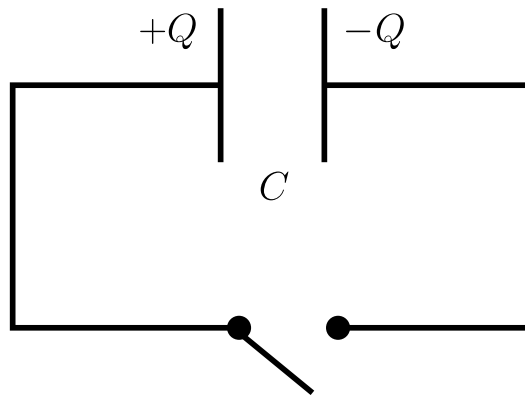
However, if we provide a conductive path between the two plates for some short period of time by closing the switch:
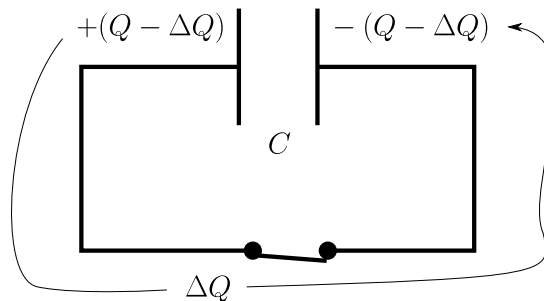
then some positive charge ΔQ travels from the positively charged plate to the negative plate[1]. Now when we open the switch back up:

$$+(Q - \Delta Q) \qquad -(Q - \Delta Q)$$

$$C$$
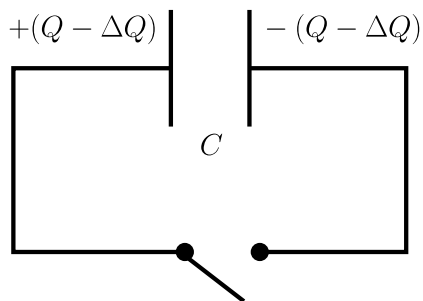
we find the charge on the positive plate is reduced by an amount ΔQ, and likewise the charge on the negative plate is increased by ΔQ, so that the potential difference V' between the two plates now is reduced by an amount $\Delta V = \Delta Q/C$.

In this sense we would say that a lone capacitor is not a source of constant voltage. What we would like is the ability to conduct electric charge from one point to another without reducing the potential between those points. This is accomplished by using a battery. In the following we describe the first and possibly simplest battery: the voltaic cell.

To begin with we take a basin of water:

H_2O
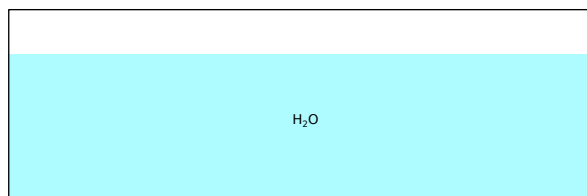
and add to it some concentrated sulphuric acid:

H_2SO_4

H_2O

---
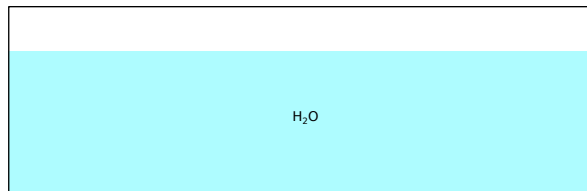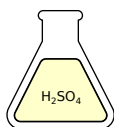
[1] As we have discussed, in metals it is actually the negatively charge electrons which conduct current, so really we should consider a negative charge $-\Delta Q$ leaving the negative plate and settling on the positive plate. Either way we arrive at the same conclusion.

In its concentrated form, sulphuric acid contains almost entirely intact H_2SO_4 molecules[2]. When diluted with water, however, the H_2SO_4 will almost completely dissociate into hydrogen sulfate (HSO_4^-) and a proton (H^+), and from there about 5% of the HSO_4^- molecules will themselves dissociate into a sulfate ion (SO_4^{2-}) and proton:
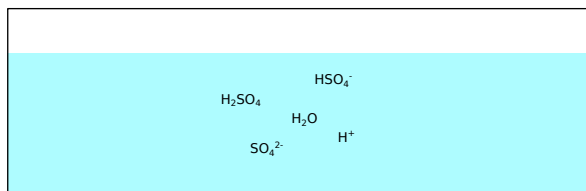
In other words, the reaction

$$H_2SO_{4(aq)} \longleftrightarrow SO_4{}^{2-}{}_{(aq)} + 2\,H_{(aq)}{}^+ \tag{1}$$
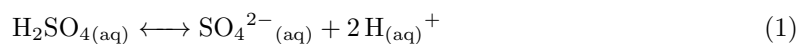
starts out badly unbalanced, since there initially no sulfate ions in the solution. As a result, the sulphuric acid will dissociate until the reaction reaches equilibrium. This occurs when the rate at which the sulphuric acid dissociates equals the rate at which a sulfate ion and two protons meet to form sulphuric acid[3]

Once reaction (1) reaches equilibrium, we complicate the situation further by putting a plate of zinc (Zn) and a plate of copper (Cu) into our basin:
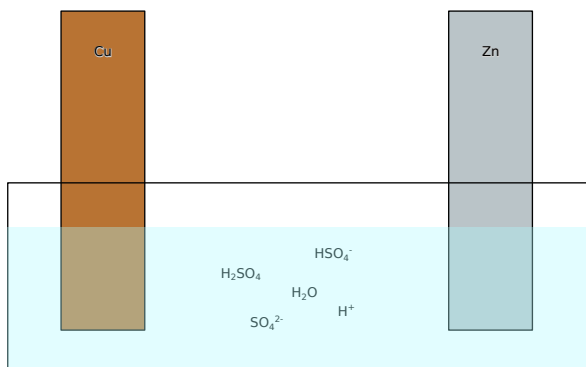
With the introduction of these materials, a number of additional reactions become possible. For one, the zinc atoms in the plate can dissolve into our solution as Zn^{2+} ions.

$$Zn_{(s)} \longleftrightarrow Zn^{2+}_{(aq)} + 2\,e^-_{(Zn)} \tag{2}$$

---

[2]This is similar to pure water, where only a tiny fraction (10^{-7}) of the H_2O molecules are dissociated into H^+ and OH^-.

[3]For the sake of simplicity we ignore in the remainder of the discussion the intermediate hydrogen sulfate ion HSO_4^-.
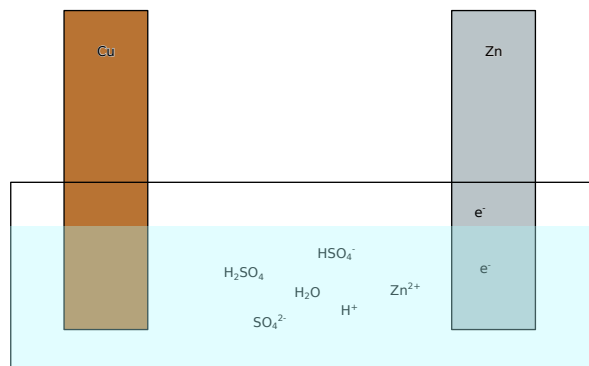
When we first introduce the zinc plate, there are no dissolved zinc ions, so again the reaction is out of equilibrium. Therefore we will expect, at least initially, some spontaneous dissolution of zinc ions into solution. Crucially we note that the dissolution of a zinc ion leaves behind a couple electrons in the zinc plate, so that as the zinc dissolves the plate acquires a negative charge.

With an appreciable concentration of dissolved zinc, we are now forced to consider the formation of zinc sulfate:

$$\text{Zn}^{2+}_{(aq)} + \text{SO}_4{}^{2-}{}_{(aq)} \longleftrightarrow \text{ZnSO}_{4(aq)} \tag{3}$$
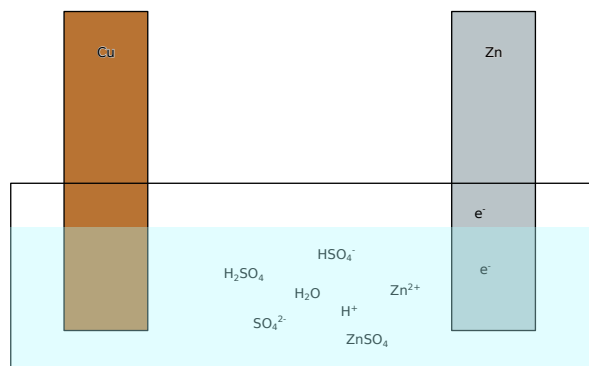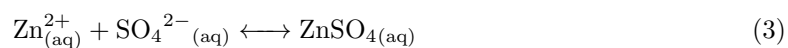
Since we begin with no zinc sulfate, the reaction will proceeds from left to right. The concentration of sulfate ions therefore decreases as they consumed in the production of zinc sulfate. This throws reaction (1), which was in equilibrium before the introduction of the plates, out of balance. To restore balance, additional H_2SO_4 is consumed, which in turn increases the concentration of H^+. The excess H^+ combines at the copper plate, forming H_2 gas which bubbles out of solution:

$$2\,\text{H}^+_{(aq)} + 2\,\text{e}^-_{(Cu)} \longleftrightarrow \text{H}_{2(g)} \tag{4}$$
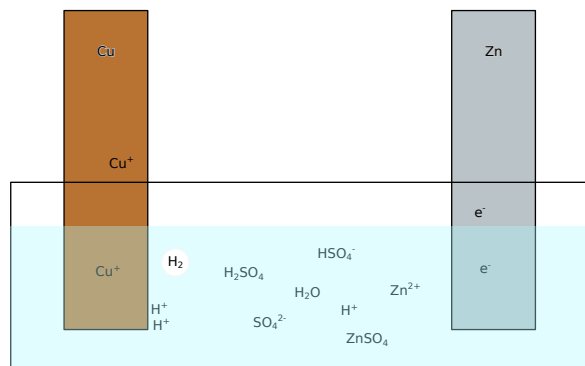
Note that this reaction results in positive charge accumulating on the copper plate. Overall we could say that the voltaic cell consumes sulphuric acid and zinc, producing zinc sulfate and hydrogen gas and transferring two electrons from the copper plate to to the zinc plate:

$$H_2SO_{4(aq)} + Zn_{(s)} + 2\,e^-_{(Cu)} \longrightarrow ZnSO_{4(aq)} + H_{2(g)} + 2\,e^-_{(Zn)} \tag{5}$$
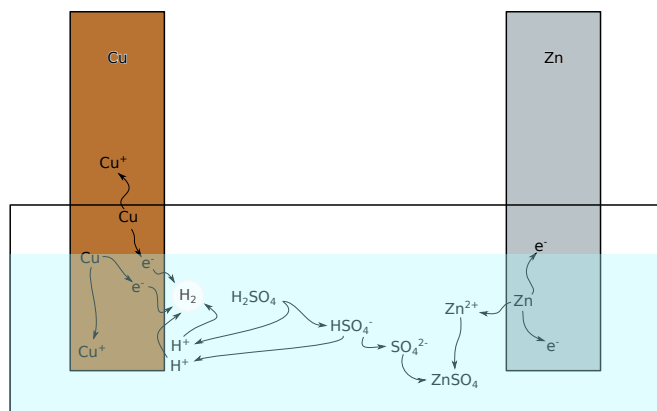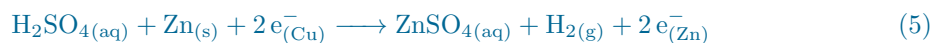
For the sake of future reference we will call reaction (5) the "**battery reaction**", to contrast it with the "**conduction reaction**", which we introduce later. When does this **battery reaction** reach equilibrium?

Qualitatively we can see that as more zinc is dissolved by the sulphuric acid, the copper and zinc plates become increasingly positively and negatively charged, respectively. This charge accumulation will act to decelerate the reaction for two reasons. First, the negative charge on the zinc plate will suppress the further dissolution of zinc, since the positive zinc ions will be attracted back onto the zinc plate. Secondly, the positive charge on the copper plate will both repel the positively charged protons and discourage the further removal of its (negatively charged) electrons. When the electric potential between the two plates becomes sufficiently large, we would expect these electrical forces to completely balance the chemical forces driving the zinc dissolution, and that the reaction

6

would therefore come to a halt. The voltage at which this occurs is known as the *emf* ("ee em eff") of the battery, or its *open circuit voltage*, for reasons we shall soon see.

The emf of a battery depends of course on the materials composing it. For instance, the Zn/Cu "Galvani cell" described here has an emf of 1.1 V, while your standard alkaline cell has an emf of 1.5 V. A 9 V battery is obtained by connecting six of these cells end-to-end (i.e. in series). A lithium ion battery has a whopping 3.6 V cell, and, unlike the previous two examples, has the benefit of being rechargeable.

Our original aim was to describe a device that produced a constant electric potential between two points even as charge was being conducted between the points. Does the battery as we've described it meet this criteria? To answer this we need to consider one further "reaction": the direct conduction of electrons between the two plates:

$$2\,e^-_{(Zn)} \longleftrightarrow 2\,e^-_{(Cu)} \tag{6}$$

This is the "**conduction reaction**" mentioned earlier. The solution inside our basin can conduct ions but not electrons, so as it stands this reaction is prohibited. We can activate this reaction by joining the two plates with a conductor, i.e. by closing the switch in the diagram below:
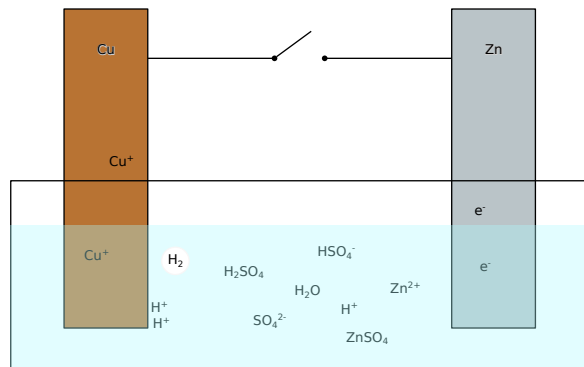
When we do this, in which direction does the **conduction reaction** proceed? From left → right or right → left? Well, an electron in the zinc plate will experience both attraction to the Cu^+ ions in the copper plate and repulsion from the excess electrons in the zinc plate. Given the opportunity then, electrons will spontaneously conduct from the zinc plate to the copper plate, i.e. the **conduction reaction** proceeds from left to right:

7
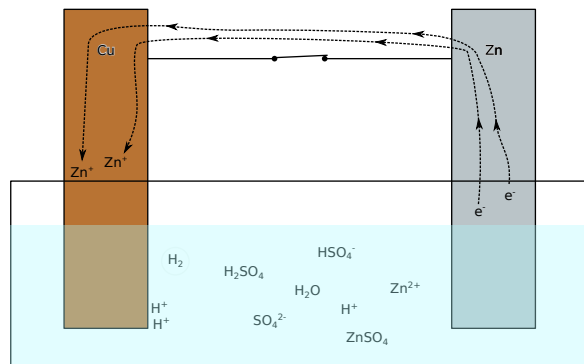
This reaction reaches equilibrium when there is no longer any electrical force driving the reaction. This occurs when electric potential on both plates very nearly the same, or, in other words, when the charge built up on the plates by the **battery reaction** is completely undone and the plates are returned to their original neutral state. What occurs when we now open the switch, deactivating the **conduction reaction**?

The key observation is that we are now essentially where we started when we first introduced the plates into the solution[4]. The discharging of the plates eliminates the electric forces, allowing the dissolution of zinc and accompanying charging of the plates to proceed all over again[5]:
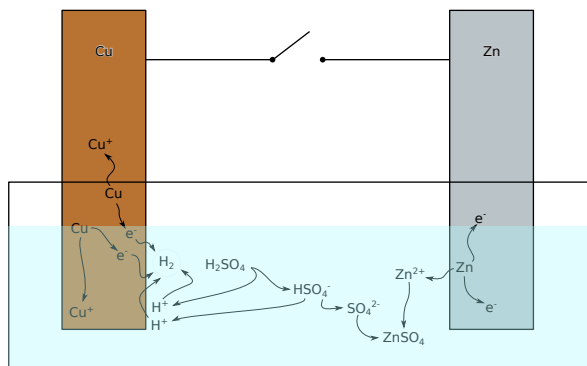
We see then that, in contrast with the simple capacitor, the chemistry of the battery will act to reestablish the full emf at the battery's plates, even after being discharged.

It is however still not clear is whether our battery can maintain the its full emf at its plates while the **conduction reaction** is activated, i.e. *while* charge is conducting between the plates. Indeed when we connect a conductor between the plates we observed in class (and also reasoned above) that the electric potential between the plates goes to zero. In this case any electron that is transferred to the zinc plate from the copper plate by the **battery reaction** is immediately conducted back to the copper plate by the **conduction reaction**. In chemistry terms we would say that the **battery reaction** is the *rate-limiting reaction*.

If somehow we are able to arrange so that electrons can conduct between the plates but only conduct *slowly*, then we can reverse this state of affairs so the **conduction reaction** is now the rate limiting reaction. In this case then the discharging of the plates via **conduction reaction** takes much longer than the charging of the plates via the **battery reaction**. In other words, any electron transferred from the zinc plate to the copper plate via the **conduction reaction** is immediately replenished by the charging accomplished by the **battery reaction**. The result in this situation

---

[4]To be fair, the battery consumed some sulphuric acid and dissolved some zinc in the process of charging the plates, so the system is not exactly how it was when we started. A typical battery, however, will contain a sufficiently high initial concentration of acid so that the battery can charge and discharge many ($\sim 10^6$ for a typical AA battery) times before the acid's concentration is significantly diminished. The key is that equilibrium in the battery is established primarily by the change in the electrical potential difference between the plates, not by significant modification of the reactant concentrations.

[5]In a more technical sense we can observe that with the discharging of the plates we have thrown the **battery reaction** way out equilbrium, since we greatly increased and decreased respectively the concentrations of electrons on the copper and zinc plates from their values when the **battery reaction** was in equilibrium and we had the full emf of the battery between the two plates. Therefore, to reestablish equilbrium we require the **battery reaction** to proceed again from left to right, as it did when we first introduced the plates.

is that charge is permitted to conduct betwen the plates without the potential between them ever deviating significantly from the full emf of the battery.

3   Resistors, Current, and Resistance

3.1   Drift speed v_d

An object that facilitates this constriction of charge flow is called a *resistor*. It can be thought of as a very poor conductor. How is it that a resistor obstructs the flow of charge? Perhaps unsurprisingly, there are multiple factors that go into determining whether or not an object will be a good resistor, and over the rest of the lecture we will touch upon all of the main ones.

However, there is one property common to all resistors which conceptually gives rise their defining behavior. In any resistor, we find that the charge carriers are subjected to frequent collisions with their surrounding environment. These collisions can be thought to originate from defects or imperfections with the microscopic arrangement of the atoms in the material composing the resistor. Here is a diagram illustrating a few examples of such defects:
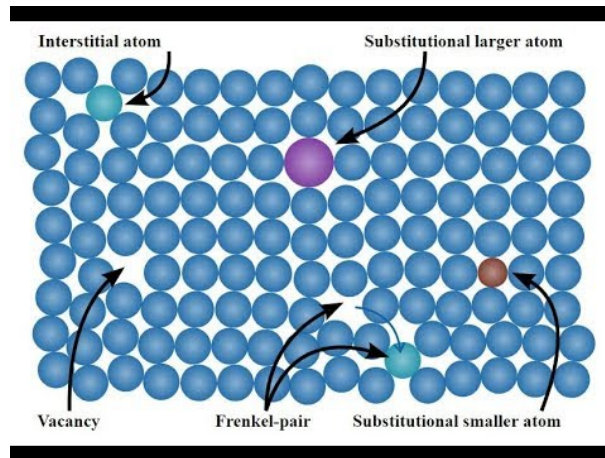
The precise nature of these defects are not relevant to us, only the fact that in any real world material these defects exist, and in materials that make good resistors these defects are abundant. If such a material existed that was totally free from defects, and we were to apply an external field in this material, its charge carriers would accelerate indefinitely since they would be at all times experiencing a force $F = qE$, where E is the strength of the electric field and q is the charge of the charge carrier. In a real world material, this acceleration could only occur for so long before a charge carrier encounters a defect and scatters from, resulting in a randomization of its velocity:

charge carrier velocity and trajectory
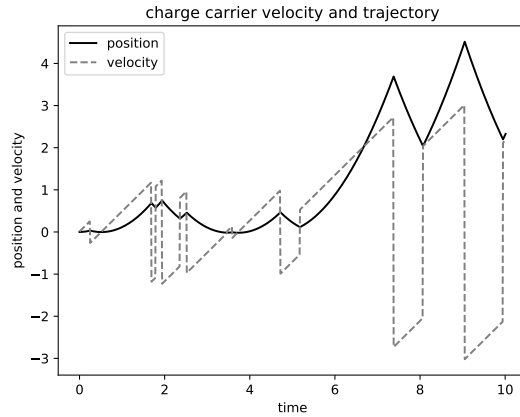
In the graph above we plot a one dimensional example of this, where we have segments of uniform acceleration interrupted by "kinks" where the charge carrier's velocity is reversed as a result of a collision with a defect.

Let τ be the average time a charge carrier spends between collisions. On time scales short compared to the τ the charge carrier experiences uniform acceleration, but on time scales somewhat larger than τ its motion appears fairly random as a result of frequent collisions. How does its motion appear to behave on time scales that are much larger than τ? Well, because of the constant force qE being exerted at all times on the charge carrier we can say that, on average, the particle experiences an increase $v_d \equiv a\tau$ in velocity between collisions due to the constant acceleration $a = \frac{F}{m} = \frac{qE}{m}$. We might suspect then that its motion of very long times scales is a constant "drift" in the direction parallel to the electric field:

charge carrier velocity and trajectory
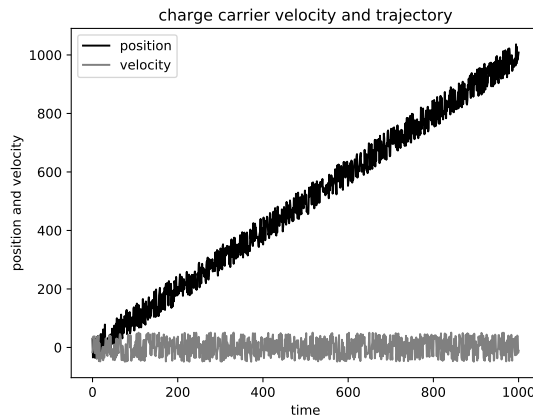
so that the charge carrier moves on average with a constant "drift velocity":

$$v_d = \frac{qE}{m}\tau \tag{7}$$

The fewer defects present in the resistor, the longer the time τ between collisions, and the larger

10

the resulting drift velocity v_d. We close the main discussion of drift velocity with a list of some mean collision times τ for some different charge carriers in different materials.

charge carrier	material	mean collision time (fs=10^{-15} s)	mean free path ($\text{Å} = 10^{-10}$ m)
electron (e^-)	silicon	1000	1000
electron (e^-)	copper	30	30
sodium ion (Na^+)	liquid water	14	.06

We are able to grow silicon crystals of extremely high purity, so that electrons can travel for relatively long distances (about 400 silicon atoms) before encountering an obstacle. Electrons in copper do not get nearly as far (about 15 atoms) before colliding, and sodium ions in liquid water will collide multiple times before getting past even a single water molecule.

Later we will find that copper, despite having a much shorter time τ, is a much better conductor than pure silicon. This is because the number of *charge carrying* electrons per unit volume, i.e. the number of electrons that are not *bound* to any nucleus, is much higher for copper than for pure silicon.

3.1.1 Analogy: terminal velocity

This situation where a constant force gives rise to a constant average velocity occurs not only for microscopic objects but for everyday objects as well. A person falling through the sky with their parachute deployed certainly does not continue accelerating at a constant rate of $9.8\,\mathrm{m\,s^{-2}}$. In fact a skydiver will reach a constant velocity, known as their *terminal velocity*, even before they deploy their parachute[6].

This occurs as a result of a enormous number of collisions occuring constantly between the skydiver and the gas particles in the air. Because of the skydiver's mass is vastly larger than an air particle's, each collision results in an extremely small randomization of the skydiver's velocity, and so a relatively long time is required for the skydiver to reach terminal velocity.

Charge carriers, on the other hand, are either have masses that are much lighter than the atoms they collide with (when they are electrons), or have masses that are of the same order of magnitude (when they are ions). As a result, charge carriers reach "terminal velocity" over the course of a few collisions. Therefore, we can assume that, to a very good approximation, a charge carrier in an electric field of strength E moves at *all* times with an average speed $v_d = \frac{qE}{m}\tau$.

3.2 Current

In the previous section we established that charge carriers in a resistor move on average with a constant velocity that is proportional to whatever electric field E exists in the medium. A externally applied electric field, then, will set induce a steady *flow of charge* in a resistor. This flow of charge is called an electric *current*.

To give a quantitative definition of current, we must first be more specific about what we mean by a resistor. Before we can consider some object to be a resistor, we must first specify its two ends where charge/current can enter and exit. As an example we can take a block and pick at least three different ways for current to enter and exit:

---

[6]A skydiver's terminal velocity ends up being about 150mph. If there were no air to impede his/her fall, simple energy conservation tell us they would reach a velocity of about 600mph before striking the ground.
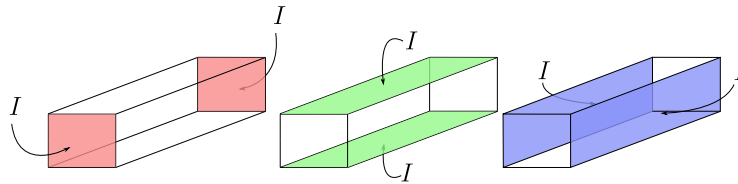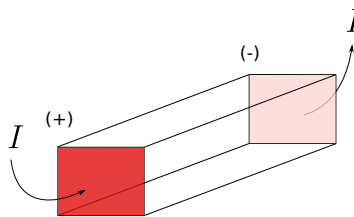
You will see from one of the homework problems that an object's ability to restrict the flow of charge (i.e. its *resistance*, which we will define more precisely later) will depend on which two ends we choose to conduct current through.

Having selected the two ends, say these two:
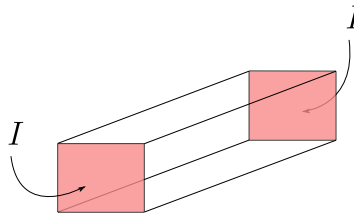
we now pick one of these two, say the front (dark shaded) end, to be the positive (+) end:

and the other end, which we've picked to be the light shaded end, to be the negative (-) end.

Having made these selections, we can now define the current through the resistor. We can state this in four equivalent ways:

The current, denoted I, is the rate at which:

    (a) positive charge enters the positive end,

    (b) positive charge exits the negative end,

    (c) negative charge exits the positive end, or

    (d) negative charge enters the negative end.

Current is evidently a charge divided by a time. Indeed, the SI unit of current, which is called an *ampere*[7] and denoted by the symbol A, and is defined so that $1\,\mathrm{A} = 1\,\mathrm{C\,s^{-1}}$.

To get a feel for what are large and small amounts of electric current, here is a list of some currents:

---

[7]or *amp* for short

- Current supplied by car battery to start a car in the winter: $300\,\text{A}$

- Current through glowing incandescent light bulb: $0.5\,\text{A}$

- Current supplied by watch battery (lasts about 3 years): $1\,\text{µA}$

- Current in a lightning bolt: $30\,000\,\text{A}$

3.3   Resistance

So we have learned that resistors are two-ended objects that are composed of a material containing of charge carriers that travel with some drift velocity $v_d = \frac{qE}{m}$ in response to an electric field. Knowing this, can we determine a relationship between the *voltage* V applied across the two ends of our resistor and the current I that is conducted through the resistor in response to this voltage?

We will work out this problem for the case of our rectangular resistor introduced in the previous section, but the result we obtain (Ohm's law) holds in general. To begin, let's use the symbol A for the area of the resistors end faces A and the symbol l for its length:
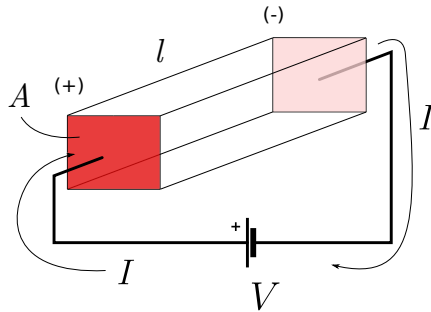
To determine the current, we ask ourselves, "In a time Δt how much charge ΔQ exits the negative end of the resistor?" Once we figure this out, we calculate the current I by taking the ratio $\Delta Q/\Delta t$.

To determine ΔQ we take the following steps:

1. The first thing we notice is that between the two ends of the resistor the electric field is constant and given by $E = \frac{V}{l}$.

2. Therefore the charge carriers in the resistor are all moving with a velocity $v_d = \frac{qE}{m}\tau$.

3. In a time Δt, these charge carriers will travel a distance $\Delta x = v_d \Delta t$.

4. From this we conclude that any charge carrier within a distance Δx of the negative end of the resistor will leave the resistor within the time interval Δt:
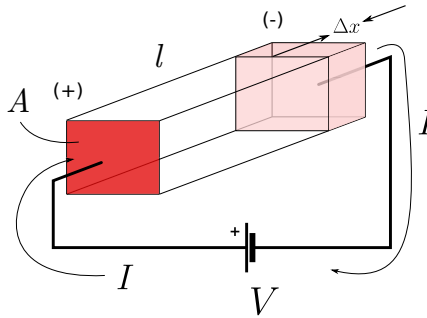
5. Therefore, the charge ΔQ exiting the negative end of the resistor is equal to the charge contained in the light shaded volume in the diagram above. This is given by $A\Delta xnq$, which is the product of:

- The volume $A\Delta x$ of the region,
- the density (number per unit volume) n of charge carriers in the resistor, and
- the charge q of each charge carrier

Putting this all together we get:

$$
\begin{aligned}
I &= \frac{\Delta Q}{\Delta t} \\
&= \frac{A\Delta xnq}{\Delta t} \\
&= \frac{A\Delta tv_dnq}{\Delta t} \\
&= Anq(qE\tau/m) \\
&= \frac{nq^2\tau A}{ml}V
\end{aligned}
\tag{8}
$$

We arrive at the most important result of the section:

The current I through a resistor is proportional to the voltage V across it.

Conventionally we express this relationship in the following form, which we call *Ohm's Law*:

$$V = IR \tag{9}$$

The constant R as you may have guessed already is the *resistance* of the resistor. From the above equation we can see it is the ratio of a voltage and a current. Accordingly the SI unit of resistance is $1\,\text{V}\,\text{A}^{-1}$ which is known as given the symbol Ω, called an *ohm*.

3.4 Resistivity

From equations (8) and (9) we find that the macroscopic resistance R of a resistor is related to its microscopic properties by the following equation

$$R = \frac{m}{nq^2\tau} \cdot \frac{l}{A} \tag{10}$$

14

The second factor has parameters that describe the *geometry* of the resistor. This factor tells us that a long, narrow resistor has a higher resistance than a short, fat resistor. This is consistent with our experience of water flowing through pipes, where for equal water pressures the short pipe with a large diameter will conduct a higher flow of water than a long pipe with a small diameter.

The first factor, on the other hand, has parameters that all describe the *material* composing the resistor. For convenience we often combine all the terms in the first factor into a single quantity called the *resistivity* of the material:

$$\rho = \frac{m}{nq^2\tau} \tag{11}$$

so that a resistor made of a material of resistivity ρ has a resistance R given by

$$R = \rho\frac{l}{A} \tag{12}$$

From the above equation we see that resistivity has SI units of $\Omega\,\mathrm{m}$. Copper, for instance, has a very small resistivity of $1.7 \times 10^{-6}\,\Omega\,\mathrm{cm}$, while pure silicon has a much larger resistance, somewhere in the ballpark of $100\,\Omega\,\mathrm{cm}$. The reason why copper is so *conductive*, i.e. why it has such a *low* resistivity, is because every copper atom in the resistor contributes one electron as a charge carrier. Contrast this with silicon, where all the electrons are quite strongly bound to a silicon nucleus, with the result that, at room temperature, only about one in every 10^{12} silicon atoms contributes a "conduction electron".