
Electricity & Charge

All normal matter is made up of charged particles. “Charge” is just a way to measure how strongly two objects attract (or repel) one another electrically. Charge is similar to mass in that the product of the force between two charges depends on the sizes of those charges and is inversely proportional to the square of the distance separating them, just as we saw when we studied the law of gravity. Notable differences between the two situations include the possibility of repulsion between like charges (all masses attract one another gravitationally – even antimatter) and the vastly greater strength of the electric force.

We observe charges to be quantized (i.e., not continuous, but having a smallest possible size, just as we now know water is not a continuous substance but a large number of molecules) in units of the charge on the electron, usually represented as \( e \). The SI unit of charge is called the Coulomb (C) (we’ll define it later) and it is a very large number of charges. In fact, each electron has a charge of only \(-1.6 \times 10^{-19}\) C. That means it would take a few billion billion of them to make one Coulomb. By convention, the charge on the electron is defined as negative, meaning the charge on the proton is positive. The signs are opposite, but the charges are exactly the same magnitude, to the highest accuracy we can measure.

As always, there will be a few enhancements and corrections to the statements above that we’ll examine later. One is that experiments tell us that we can actually have charges which are 1/3 and 2/3 the size of the basic charge \( e \) when we look at the particles which make up protons and neutrons (called quarks). These have never been isolated individually, but their existence has been demonstrated indirectly. For this reason, we’ll still consider \( e \) to be the fundamental unit of charge.

You’ve probably known since childhood that like charges repel and unlike charges attract. To this idea we’ll add the fact that charge is conserved. This law is absolute. No experiment has ever shown the overall destruction or creation of charge. We can separate charges very easily – rubbing a glass rod with a piece of fur, or your hair with a balloon will demonstrate this. These are methods of mechanically transferring electrons from one place to another. When you do this, the material that gave up the electrons becomes positively charged (having lost some amount of negative charge) and the material that picks up the electrons is negatively charged. The sum of the charges, though, is still zero (assuming the objects were neutral in the first place).

Conductors & Insulators

Not surprisingly, certain materials conduct electricity better than others (usually, we see that good conductors of heat are also good conductors of electricity). In a good conductor, the outermost electrons of the atoms (at least some of them) are not tightly
bound. They are essentially free to roam around the conductor, and when they move, they carry charge from one place to another (generating a current, which we’ll talk much more about in a few chapters), or conduct electricity. Among the best conductors are gold, silver, copper, and aluminum (they’re also very shiny, and that’s not a coincidence). Things like quartz, glass, rubber, most plastics, wood, etc. are not good conductors, and are therefore called insulators. The atoms in these materials hold their electrons very tightly.

**Charging Objects**

If we put an excess of electrons on an object, we say that it has been charged. We could also take electrons from the object, and put a positive charge on it. If we want to try this on a conducting sphere, it needs to be insulated from the Earth. The reason is that the Earth is a huge conductor, and will gladly grab all of the excess charge it can reach. This is why we use lightning rods to direct lightning to the Earth – it can soak up the charge easily.

We can charge the object by touching it with a charged rod and letting some of the charge transfer (it’s easier to say that than to say “letting electrons move from the object to the rod if the rod is positively charged” or “letting electrons move to the object from the rod if the rod is negatively charged”). The excess charge will quickly (very quickly) move to the surface of the metal sphere. It will do this because that’s the way it can get as far away as possible from the other charges of the same sign that came over with it from the rod. This is charging by contact.

We can also charge by induction. Imagine connecting the metal sphere to the Earth (“grounding” it) and bringing a negatively charged rod nearby. The electrons in the sphere near the rod are pushed away, and the grounding wire gives them an exit to use. Some of them leave the sphere entirely, and will only return when the charged rod leaves. What if we break the grounding wire connection before letting the rod leave? Now we’ve driven electrons into the Earth, and they can’t get back to the sphere, so it is left with an overall negative charge. (Notice that it’s exactly the same size as the opposite charge we’ve added to the Earth, so we’re still not creating or destroying charge, just moving it).

We couldn’t do the same trick with an insulating sphere, because the electrons near the rod aren’t free to move large distances away from their atoms, much less leave entirely. All they can do is shift position a small amount (small on the atomic scale) and spend more time on the far side of the atom they’re bound to. That will tend to make the insulating sphere’s surface slightly positive and give a small attraction between it and the rod.
Current

If charges move, they will create an electric current. Current is just the rate at which the charge in a region changes. In an infinitesimally small time, we will see that the current $I$ is found from

$$I = \frac{dQ}{dt}$$

The units of current are Coulombs/second, or amperes. Whenever a current exists, we have moved away from electrostatics and into electrodynamics.

While electrons generally do the moving (and therefore constitute the current) when there is a potential difference applied to something, it was originally believed that positive charges were moving to cause the electric current. Electrons are said to have a negative charge, though, so the flow of electrons is from a negative battery terminal to a positive battery terminal. When we look at currents, we'll generally refer to them as moving through a circuit from the positive terminal to the negative terminal, even though we know that's not exactly what's happening.

When we place a battery in a circuit, it will push electrons from its negative end through the circuit (doing some work, in general) and back to the positive end. This flow of charges is called the current and is measured in coulombs per second, or amperes. For example, your car battery may be able to deliver several hundred amperes of current to your starter. That means that if you picked a point somewhere along the battery cable and could count charges, you’d see hundreds of coulombs of charge go by every second.

A battery produces direct current (dc), which means the current is always in the same direction (from the positive pole to the negative pole – this is called the conventional current, and the idea dates from a time when people thought positive charges were doing the moving. Today we know that it’s generally the negatively-charged electron which does the traveling, and the flow of electrons is therefore opposite to the direction of the current). The wall outlets provide alternating current (ac) in which the electrons are pushed back and forth, changing their direction many times per second. The difficulty or ease with which electrons are set in motion by a potential is described by the resistance of the material containing the electrons.

Resistance and Resistivity

The resistance of a given object depends on a few factors: first, the composition of the object. Wires are made of copper, not glass. This measurement of the inherent ability of a material to conduct electricity is called resistivity and is usually abbreviated by the letter $\rho$. For two objects made of the same material but which have different sizes or
shapes, resistance will also be different. Resistance is proportional to the length of the wire (or object) and inversely proportional to its area. In other words, a fatter pipe lowers the resistance, and a longer pipe increases it. The formula for resistance can be written as

\[ R = \frac{\rho L}{A} \]

where \( L \) is the length of the current path through the object and \( A \) is the area of the wire. This tells us that the units of resistivity must be ohms per meter. From the table in your book, you can see that some metals like copper and silver have resistivities lower than \( 2 \times 10^{-8} \, \Omega \cdot m \). At the insulating end of the spectrum, resistivities are in the neighborhood of \( 10^{14} \, \Omega \cdot m \) or higher. This is why wire is typically made of copper (good conductor) covered with plastic (very bad conductor). If the cord for your lamp was not insulated, the two wires would eventually bump into each other and provide a way for the charges to return to the wall without lighting your lamp (short circuit). If you happened to step on both wires, current would go through your foot!

The dependence of resistance on area can be seen when we look at the sizes of wire used in various situations. The wires inside your radio or computer are typically carrying very small currents and therefore can be very small themselves. The main power cable into your house is probably about the size of your wrist since it’s carrying more current. This is why you sometimes hear that extension cords can be dangerous. If you’re using a thin cord (high gauge), you’re putting a resistance in between the power and the device. You can usually feel the cord getting warm. Longer cords should also be fatter to counteract this problem.

What causes resistance? It’s collisions between the moving electrons and the rest of the wire (including impurities in the wire). Imagine rolling bowling balls down hills covered with tree stumps. The balls will roll a short distance and then collide with a stump, losing speed and possibly going back uphill briefly. They’ll get downhill eventually, but it will depend on the spacing of the tree stumps. Current in a wire is also sometimes compared to water in a hose. An interesting similarity between the two is that, in both cases, the speed of the particle (water or an electron) is very much less than the speed at which the signal travels. As soon as you flick a light switch, the electricity travels with a speed not far from that of light. The individual electrons have speeds (because of all the collisions slowing them down) that are typically a few centimeters per second or so. The speed difference in a hose is also noticeable – if the hose is filled with water already, it will start spraying almost as soon as the faucet is turned on. An individual water molecule, though, takes a while to get from the faucet to the end of the hose.

The resistance of most substances is temperature-dependent. At higher temperatures, the resistance increases. For some materials, they can be cooled to a point where all resistance to the flow of current disappears. This is called superconductivity and was
discovered in the early 1900’s, but the highest temperatures at which it could be observed were in the neighborhood of 20 K. In the mid 1980’s, new kinds of superconductors were discovered which remained superconducting to around 175 K or so. This is still very cold, but it’s much easier to cool something to 175 K than 20 K! A material that is superconducting at room temperature would be a major breakthrough and would save an unbelievable amount of energy every year which is now lost to heat.

**Current and Voltage Sources**

An *ideal* voltage source is one that has a constant voltage across its terminals for *any* current delivered. A battery is not an ideal voltage source: as the load increases, the voltage typically drops. This is why we keep some of the old 12-volt batteries around in the lab. Even though they are so old that they can’t really function in devices meant for them, like large flashlights, they can still charge small capacitors to 12 V for use in other labs because the current drawn during charging falls off exponentially. The batteries can keep a 12 V potential difference between the terminals because so little is asked of them when they charge the capacitors.

An *independent voltage source* of $V$ volts will keep that potential no matter what current is drawn. The symbol for an independent voltage source is just a circle with a "+" sign and a "-" sign inside it, like this:

![Independent Voltage Source](image)

There are also *dependent* voltage sources, where the voltage supplied depends on some other circuit parameter, like the current through some other element or the voltage between two points. This is drawn the same way as the independent source, except the circle is replaced by a diamond:

![Dependent Voltage Source](image)

In the same way, we can define a *current source* to be something that always carries the same current through itself, no matter what the voltage between its two terminals may be. Again, we have *independent* and *dependent* sources, and the dependent sources (just as in the voltage case) can be controlled by either a voltage or a current elsewhere in the circuit. The symbols are the same as for voltage sources, but the plus and minus are replaced by an arrow indicating the direction of the current.

![Independent Current Source](image) ![Dependent Current Source](image)
Ohm’s Law

The flow of charge through a circuit is very much like the flow of water through pipes. The battery is like a pump that pushes water through one end of the pipes and pulls it out the other end. Just as water doesn’t flow through the pipe without friction, electric charges don’t move through ordinary conductors without some loss of energy. This loss is called resistance for the obvious reason that the movement of charges is being resisted by the material they’re moving through. For many materials, the resistance of a given size of material is a constant and we get the famous equation shown below:

\[ V = I R \]

This is known as Ohm’s Law after the person who discovered it. This says that as the potential difference across a piece of material increases, the current through that material increases in direct proportion. If we connect the two ends of a 1.5 V battery to a black box (we don’t know anything about what’s inside – it just has two wires sticking out of it) and get a current of 5 amperes, we can then connect it to a 9 V battery and know that we’ll get a current of 30 amperes if the box obeys Ohm’s law. Not everything obeys Ohm’s law!

The units of resistance are volts per ampere (V/A) or ohms. The common symbol for the ohm is the Greek capital omega, or \( \Omega \). When we’re selecting wire for a circuit, we generally want it to have the lowest resistance possible (and we’ll typically assume that the resistance of the wire is zero unless we have a particular reason to worry about these tiny resistances). What would our equation above predict for the current produced if we connect the poles of our 9 V battery with a wire of zero resistance? The equation would predict an infinite (or more accurately, undefined) current. What stops this from happening is the fact that the battery has internal resistance of its own, so we don’t have to worry about the math failing. In reality, the internal resistance is usually pretty small, so you can get currents that are very large by “shorting out” the battery. Doing this with a car battery is almost guaranteed to injure someone.

Electric Power

We can find the power consumed in a circuit ( = the power delivered by a battery) by taking advantage of what we know about energy. We know that moving a charge \( q \) through a potential difference \( V \) requires an energy equal to \( qV \). We also know that energy per time is power. We can combine these to get

\[ P = \frac{dq}{dt} V = IV \]

The unit of power is still what it was before – 1 watt = 1 J/sec or, in this case, 1 Volt-Ampere. If the device we’re looking at obeys Ohm’s law, we could also write:
Using this, what is the resistance of one of the heating elements in an electric water heater if it consumes 2500 W at 220 V? It’s about 20 Ω. That means the current must be around 11.2 A.

**EMF**

If we want to talk about charges moving, rather than sitting still, they need a reason to move. That reason is called the **electromotive force** or **emf**. Emf is not really a force, though – it’s just a potential difference between two places (voltage) which, as we know, will cause charges to move in response to it. The symbol for emf is a script capital “E” or E. If we use a battery to provide the emf, it works by means of a chemical reaction that causes electrons to collect at one end and be depleted from the other end. The potential difference between the ends, or emf, is a measure of how badly the electrons “want” to get back together with the ions. If you connect the ends of the battery with a wire so that the electrons have a way to do it, many of them will move along the wire to rejoin ions (don’t do this – it will drain your battery, and for reasons we’ll see later, it will also make the wire dangerously hot). For our purposes, a battery is just a source of emf.

**Circuits**

Combining resistors with other electrically useful elements will give us an electrical **circuit**. The name is significant because there must be a complete path allowing the electrons to travel from a source of emf (a battery, for example) through the various elements and then back to the other end of the emf source. If that path is **not** complete, we have an **open circuit** and no work will be done. When we’re designing these circuits, we find that it’s easier to draw a picture of the circuit than it is to list every element by name. Right now, our circuits will be pretty simple. See the table below for pictures.

<table>
<thead>
<tr>
<th>Element</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Battery</td>
<td>![Battery Symbol]</td>
</tr>
<tr>
<td>Capacitor</td>
<td>![Capacitor Symbol]</td>
</tr>
<tr>
<td>Resistor</td>
<td>![Resistor Symbol]</td>
</tr>
<tr>
<td>Straight Wire</td>
<td>![Straight Wire Symbol]</td>
</tr>
<tr>
<td>(zero resistance)</td>
<td></td>
</tr>
</tbody>
</table>
Kirchoff’s Rules

For complicated circuits, which may have more than one battery or generator and multiple current paths, we need to use a couple of guiding principles to figure out what’s really going on. These are called Kirchoff’s rules, and the reasoning behind them will be easy to see. First, the node (or junction, or current) rule or law (abbreviated KCL). This says that, for a given node (place where 2 or more wires or circuit elements connect), the incoming currents must exactly match the outgoing currents. If this were not true, charge would pile up somewhere in the circuit, causing an emf which would oppose the driving emf. This is really just a statement about the conservation of charge, and we can return to our analogy of people in a building. You can be sure that the number of people entering a revolving door is equal to the number of people leaving it. Mathematically, we’ll say that the currents entering a junction will be positive and the currents leaving it will be negative. The sum of these positives and negatives is required to be zero.

Notice that the definition of a node used above means we could twist the ends of two pieces of wire together to make a longer wire and call the twisted part a node. In fact, we can imagine a point somewhere along a single piece of wire and call it a node, since it joins the left and right portions of the wire. This means that a single resistor connected across a single battery could be described as having an infinite number of nodes! Of course, that would not be very useful. We would get an infinite number of equations, but they would not be independent. A far more common and useful concept is that of an essential node, which is the joining of three circuit elements or wires. There are not an infinite number of these in our circuit. An essential node will give you current equations like $I_1 - I_2 = I_3$ which will be useful in solving circuit problems.

Kirchoff’s loop rule (or voltage rule) just tells us that energy is conserved. If we follow a charge all the way around a circuit, it will see as many potential drops (across resistors, etc.) as potential jumps (across batteries or generators). We can use our conventions from the node rule to assign potentials to the trip. If we go across a resistor in the direction that the current flows, we’ll say the potential drops by $IR$, while if we go in the direction opposite to the current, the potential increases by $IR$. Crossing a battery from – to + will give a jump in potential equal to the battery’s voltage. The sum of potential drops and potential jumps through the circuit will be zero.

We have a little freedom with this method, since we can pick the direction of the current if there are multiple sources of emf and we don’t really know which way it is going in the loop. We combine our two rules, applying the junction rule at each junction and the loop rule around every loop. We’ll get multiple equations with multiple unknowns, and we then have to solve them. If, at any point, we get a negative value for one of our currents, it just means we picked the wrong direction for it.

In a given circuit, each branch will have a current through it. In your second semester of introductory physics, you probably used Kirchoff’s laws and the branch current method
to solve a few circuits that were too complicated to solve by reduction (replacing multiple resistors by a single equivalent resistor, etc.) These two rules will allow us to solve our circuit using the branch method, but they will generate quite a few simultaneous equations for us to solve. If you’re doing this by hand, and you need to find the determinants of matrices, you’ll quickly see that fewer equations would be better.

**Internal Resistance**

As mentioned before, a battery (or a generator, for that matter) has some internal resistance of its own. This is usually very small, but because it will effectively be placed in series with the rest of the circuit, there will be a voltage drop across it. The true voltage across the terminals is called the **terminal voltage**, and will be lower than it would be without the internal resistance. As the resistance of the rest of the circuit decreases, the importance of the internal resistance increases. Higher current through that internal resistance means a higher voltage drop across it.

**Resistors in Series**

If we have more than one resistor in a circuit, one way we can combine them is to join them at one end. If we do this, all of the current must pass through each resistor. This obviously acts to increase the resistance of the circuit. There will be a drop in voltage from one side of a resistor to the other. The size of the drop is determined by the current flowing through the resistor and its resistance.

In the diagram above, it is clear that the current goes through each resistor. In this situation, the resistance of the two resistors is equal to the sums of their individual resistances. In other words,

\[ R_{Total} = R_1 + R_2 + R_3 + \ldots \]

for multiple resistors in series. For example, if the battery produces a potential difference of 18 V and the resistances of the two resistors are 2 Ω and 10 Ω, what will happen in the circuit? The total resistance, by our formula above, must be 2 Ω + 10 Ω = 12 Ω. That means the current must be V/R or 18 V / 12 Ω, or 1.5 A. (Think about what would happen between the resistors if the current were different in each!). For a current of 1.5 A, the voltage drop across the 2 Ω resistor will be 2 Ω times 1.5 A or 3 V. When we get to the next resistor, the voltage drop across it will be 10 Ω times 1.5 A or 15 V. Total voltage drop = 15 V + 3 V = 18 V = battery voltage, so we’re OK. If you’ve seen old Christmas lights, they were built so that one blown light bulb (the filament of which is
just a resistor) took out the entire string. That’s because they were wired in series. If you break one part of the circuit, it’s all going to stop.

**Resistors in Parallel**

If the resistors in a circuit are arranged so that there are multiple separate paths for the current to take, they are said to be in **parallel**.

![Resistors in Parallel Diagram]

In the drawing above, a given electron will pass through *either* one resistor or the other, but not both. This provides a second path for the current to take and therefore reduces the total resistance to its flow. If you imagine people leaving a crowded building, we can think of them as the charges and resistors as the doors. If we have two doors in parallel (meaning you can go through either one to get outside), more people will be able to leave per second than with only one of the doors open. Putting doors in series would correspond to having to use both doors to get out, and that will obviously let fewer people leave than if they only had to go through one door or the other.

The voltage drop across each resistor above is the same as the voltage of the battery (they’re both connected to the poles of the battery individually, if we consider the wires to have zero resistance). For different resistances, these identical voltage drops mean different currents pass through each resistor. If we return to our 18 V battery and resistors of 2 \( \Omega \) and 10 \( \Omega \), we’ll find that the current through the 2 \( \Omega \) resistor is 18 V /2 \( \Omega \) or 9 A. The current through the 10 \( \Omega \) resistor is 18 V / 10 \( \Omega \) or 1.8 A, giving a total current in the circuit of 10.8 A. If we want to replace this pair with a single resistor which gives the same current flow, it would have to be 18 V / 10.8 A or 1.67 \( \Omega \). We can find that directly by learning how to combine resistors in parallel. The equivalent resistance of a number of resistors in parallel is

\[
\frac{1}{R_{\text{Total}}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \ldots
\]

Let’s check this formula on the result we just found. We should find that 1/R\(_{\text{Total}}\) is just the sum of 1 / 2 \( \Omega \) and 1 / 10 \( \Omega \), or 6/10 \( \Omega \). \( R_{\text{Total}} \) must then be 1/ (6/10 \( \Omega \) or 10/6 \( \Omega \) = 1.67 \( \Omega \). (This, by the way, is also the way to determine the mpg a van has to get before it will use less gas on a trip than 2 (or more) cars – remember that for spring break).
To summarize, resistors in series all carry the same current, but have voltage drops that depend on their individual resistances. Resistors in parallel are all at the same voltage, but carry a current that depends on their individual resistances. The resistance of resistors in series is always *greater* than the *largest* individual resistance. The resistance of resistors in parallel is always *smaller* than the *smallest* individual resistance.

Now we see why short circuits are dangerous. If the current can go through a wire laid across the battery terminals, the resistance in the wire is very low – close to zero. The circuit has the same resistance it always had. The current now has two choices. It can take a low resistance path, or a high resistance path. It will take the low resistance path and *lots* of current will flow. We know that the power lost to heat in this wire (because it has some resistance, even if it is small) will be $V^2/R$ or $I^2R$, which will be very large when $R$ is small.

How are the outlets in your house wired? They must be in parallel. If they were in series, all of your lights and appliances would go off when a light bulb blew. This also explains why plugging too many appliances into a circuit will cause your circuit breaker to trip. Each device you add in parallel lowers the total resistance of the circuit and therefore increases the current. When the current reaches a certain value, the circuit breaker will trip and the power will be cut off. This used to be done by fuses (and is still done that way in most cars). The idea behind the fuse is to isolate a small piece of wire made of a material that melts at a low temperature. If that little isolated piece of wire burns through, the circuit is open. It’s much better for that little piece of safely-contained wire to melt than for a random part of the wire inside your wall to do it!

Many of the resistive circuits we’ll eventually work with can be reduced to collections of resistors in series and parallel. Whenever you’re trying to decide whether two resistors are in series or parallel, just look at the current path: if the current has to go through both, they’re in series. If it’s required to pick a path, they’re in parallel.

**Current and Voltage Measurements**

How can we check ourselves and see if the numbers we calculate for things like currents and voltage drops are accurate? As we’ll see later, a current both responds to and creates its own magnetic field. By arranging a coil of wire between magnets, we can watch it move as a current passes through it. If we put a spring on it to resist its movement and a needle on the coil to track the movement, we’ve made a device called a *galvanometer*, which is the main part of a current measuring instrument called an *ammeter*. The other important part is a shunt resistor, which gives the current an alternate path through the device and thereby diverts large currents (most of them, at least) around the sensitive galvanometer. If the galvanometer needle deflects enough to measure small currents, we can see that very large currents would cause the needle to go around many times (or, more likely, would burn out the thin wires inside the galvanometer). As long as we know the relative resistances of the galvanometer and the shunt resistor, we can figure out the total current.
The ammeter has to be inside the circuit to measure the current since it can only measure what flows through it. That means it will be in **series** with the part of the circuit it’s trying to measure. That also means that its presence will increase the total resistance and therefore lower the current. The process of measurement alters the true value we’re trying to measure! This is not really unfamiliar to us, though. If we want to know the temperature of a small cup of coffee and we stick a large, room-temperature thermometer in it, we know from the study of heat that the coffee will cool a little as it warms the thermometer up. We minimize this effect in the specific heat lab by using quantities of water that are large compared to the mass of the thermometer. We can minimize the effect of our ammeter by making its internal resistance as small as possible.

If we connect the galvanometer in parallel with part of the circuit, we will be providing an alternate path for the current. If the galvanometer has a tiny resistance, we’ll siphon off most of the current and burn out our ammeter. To get around this, we put a large resistor in series with the galvanometer to make the current path less attractive to the charges in the main circuit. The current that does pass through the galvanometer (deflecting the needle) can be multiplied by the total resistance of the galvanometer + large series resistor to get the voltage between two points. Now we’ve made a **voltmeter**. Because the addition of a voltmeter gives a new current path, it will alter the behavior of the circuit unless the resistance in the new path is very large. Therefore, we’d like a voltmeter to have a very high resistance, just like we want our ammeter to have very low resistance.

Remember that ammeters are wired in **series** with the circuit, and voltmeters are wired in **parallel** with it. The ammeter can be thought of as replacing a tiny segment of the wire in the circuit. The voltmeter is connected between two different points to measure the potential difference. If we do this backwards, we’ll have a problem. A voltmeter in series with a circuit will make the total resistance huge and will probably stop the current. An ammeter in parallel with a circuit will (because of its low resistance) suddenly find itself routing a huge current between two points. For this reason, most ammeters have fuses that are designed to burn out first before the galvanometer can be destroyed.

**Resistance and the Wheatstone Bridge**

Very precise measurements of resistance can be made when you already have some high-precision resistors (the common resistor has a tolerance of 10%, meaning a 1 kΩ resistor could really be off by 100 Ω). Resistors $R_1$, $R_2$, and a variable resistor $R_v$ are combined with the unknown resistance $R_\text{?}$ as shown below:
Notice that this circuit can be described as two parallel branches with two resistors connected in series in each branch. The potential difference between the midpoints of each parallel branch (the middle left and middle right corners of the resistor square) will be zero if:

\[ R_j = \frac{R_2}{R_1} R_V \]

A sensitive ammeter is connected between points \( a \) and \( b \) and when it reads zero current, the equation above is satisfied.

**Delta and Wye Circuits**

As you have already seen in your first work with circuits, circuit simplification can frequently save a great deal of work. You may also have found cases where it wouldn’t work. We can expand the applicability of that technique a little by the use of what is known as a “\( \Delta \) to \( Y \)” transformation. For example, if we take the Wheatstone bridge and replace the sensitive ammeter with another resistor \( R_5 \), it seems that the circuit can’t be simplified.

Notice that the top half of the five-resistor circuit looks like the Greek letter delta (\( \Delta \)). We can break one of connections between the legs of the triangle and insert a straight
wire to make a square if we want, like this:

If we look at this upside-down, we get something that looks kind of like the Greek letter $\pi$:

The $\Delta$ and $\pi$ configurations are therefore electrically identical. What about a different way to connect three resistors? This is known as a “Y” arrangement, which is electrically identical to a “T” arrangement if you just bend $R_1$ and $R_2$ down a little.

The purpose of showing these is to look at the similarities between the $\Delta$ and $Y$ configurations as seen from outside the circuit (i.e., the three terminals which each arrangement has, and which we will call lowercase a, b, and c). These terminals are the free ends of the $Y$ and the vertices of the $\Delta$. Imagine that one of these circuits is in a box with three terminals. Is there a way to tell the $\Delta$ from the $Y$?

It turns out that there isn’t. We can design a $Y$ circuit that is identical (outside the box) to some $\Delta$ circuit, and vice versa. First, we show our initial conditions:
The transformation to go from a $\Delta$ to a $Y$ is

\[
R_1 = \frac{R_b R_c}{R_a + R_b + R_c}
\]
\[
R_2 = \frac{R_c R_a}{R_a + R_b + R_c}
\]
\[
R_3 = \frac{R_a R_b}{R_a + R_b + R_c}
\]

while the $Y$ to $\Delta$ transformation is

\[
R_a = \frac{R_1 R_2 + R_2 R_3 + R_1 R_3}{R_1}
\]
\[
R_b = \frac{R_1 R_2 + R_2 R_3 + R_1 R_3}{R_2}
\]
\[
R_c = \frac{R_1 R_2 + R_2 R_3 + R_1 R_3}{R_3}
\]

Your book shows an example of the utility of this technique in some cases.
**Nodes, Branches, Meshes, & Loops**

At this point, we need to more carefully define the terms we’ll use in solving more complicated circuits. From table 4.1 in your book, we introduce the idea of a **branch** as a path that connects two nodes and an **essential branch** as connecting two (and only two) essential nodes. A **loop** is a closed path where the starting and ending nodes are the same, and a loop that does not enclose any smaller loops in it is also called a **mesh**.

To recap, in the **branch method** we identify every branch and assign a variable $I_x$ to the current through it. We also must pick a direction for the current, which we can do randomly. A negative answer for $I_x$ at the end of the process just means we picked the wrong direction. If a circuit has $b$ branches, we need $b$ independent equations to find all of the currents. We can get $n-1$ equations using the KCL on the nodes, and we can then apply the KVL to $b - (n-1)$ loops so that we end up with a total of $b$ equations. Once we get our equations, the problem becomes one of linear algebra rather than physics, and a computer can do it more quickly than a person.

There are a few other ways to find the currents in a circuit that are (in some cases) quicker than the branch method by generating fewer simultaneous equations. In the **node voltage method**, an equation is written for each essential node in the circuit. Since voltage only makes sense when defined between two points, one node in the circuit is chosen to be the **reference node** from which all other node voltages are measured. The reference node is indicated by an inverted solid triangle ▼. The equations are based (not surprisingly) on the KCL. Since the sum of all currents leaving a node must equal zero, the equation for each node says exactly that. To find the currents, Ohm’s law is used. For example, if we are looking at essential node 2 which has connections to essential nodes 1, 3, and 4, our equation will be

$$I_{2,1} + I_{2,3} + I_{2,4} = 0$$

Each current is found from Ohm’s law as

$$I_{x,y} = \frac{V_x - V_y}{R_{x,y}}$$

where both $V_x$ and $V_y$ are measured from the reference node. You don’t use this method on the reference node itself – just the others. When the voltages are found for each node, you can use Ohm’s law to find the currents through the branches. For example, look at the (unbalanced) Wheatstone bridge below.
There is an essential node at each corner of the square. Notice that nodes \( a \) and \( b \) are only essential because of the presence of \( R_5 \); if it were missing, this would be a very simple circuit to analyze. We’ll pick the bottom node \( d \) as the reference node and label the other node \( c \) to start generating equations. Also, we’ll set the battery voltage to 17 V for definiteness.

There’s a simplification in this method when there is nothing but a voltage source between two nodes. As you can see, node \( c \) has to be 17 V above node \( d \) (the reference) because that’s the battery’s voltage. The currents flowing out of node \( a \) will be to nodes \( c \), \( b \), and \( d \) through resistors \( R_1 \), \( R_5 \), and \( R_3 \), respectively. We can find their sizes using Ohm’s law, and we get

\[
\frac{V_a - V_c}{R_1} + \frac{V_a}{R_3} + \frac{V_a - V_b}{R_5} = 0
\]

Since \( V_c = 17 \) V, this simplifies a little to
\[
\frac{V_a - 17V}{R_1} + \frac{V_a}{R_3} + \frac{V_a - V_b}{R_5} = 0
\]

Next, node \( b \) gives us
\[
\frac{V_b - 17V}{R_2} + \frac{V_b}{R_4} + \frac{V_b - V_a}{R_5} = 0
\]

We now have two equations for two unknowns. Adding them together will get rid of the last term in each, and leave us with
\[
\frac{V_a - 17V}{R_1} + \frac{V_a}{R_3} + \frac{V_b - 17V}{R_2} + \frac{V_b}{R_4} = 0
\]

We can now solve this. Let’s use the following values for the resistances: \( R_1 = 21 \ \Omega \), \( R_2 = 19 \ \Omega \), \( R_3 = 57 \ \Omega \), \( R_4 = 11 \ \Omega \), and \( R_5 = 42 \ \Omega \). Solving the simultaneous equations gives us \( V_a = 10.946 \ \text{V} \) and \( V_b = 6.904 \ \text{V} \). We already know that \( V_d = 0 \ \text{V} \) and \( V_c = 17 \ \text{V} \), so we can find the current through any of the resistors since we know the voltages at the endpoints. The current \( I_1 \) in resistor \( R_1 \) is just \( (V_c - V_a) / R_1 = 0.2883 \ \text{A} \). This current must be directed from \( c \) to \( a \) since \( c \) is at a higher potential. Similarly, \( I_2 \) is found to be \( (V_c - V_b) / R_2 = 0.5314 \ \text{A} \). Continuing, we see that \( I_3 = (V_a - V_d) / R_3 = 0.1920 \ \text{A} \) (towards \( d \), since it is the lowest potential in the circuit) and \( I_4 = (V_b - V_d) / R_4 = 0.6276 \ \text{A} \). Finally, \( I_5 \) is directed from the higher potential of node \( a \) to the lower potential of node \( b \), and its value is \( (V_a - V_b) / R_5 = 0.0962 \ \text{A} \). Check for yourself whether Kirchoff’s current law is satisfied at each node.

**The Mesh Current Method**

Another way we can analyze a complicated circuit (related to the branch current method) is called the mesh current method. Remember that the definition of a mesh is a loop that doesn’t contain smaller loops. In the mesh current method, we assign a circulating current to each mesh. Different conventions exist for this process; some people prefer to have all currents flowing clockwise, for example, while others alternate the flow direction so that an element on the boundary between two meshes has both currents flowing through it in the same direction, like this:
Although $I_1$ is clockwise and $I_2$ is counterclockwise, they both flow down through the resistor. You can see that if they were both clockwise, $I_2$ would flow up through the resistor. Since our book uses the “everything clockwise” convention, that’s what we’ll use. For a reasonably simple circuit like the one below, we find two mesh currents:

Notice that they’re labeled $I_a$ and $I_b$ rather than $I_1$ and $I_2$. That’s not an accident – we don’t want to imply that we’re talking about the currents through resistors 1 and 2 since $R_2$ has both currents going through it.

Once you have your mesh currents in place, you can now use the KVL to get your equations. From the left loop, we would get

$$V_1 - I_a R_1 - I_a R_2 + I_b R_2 - V_2 = 0$$

Notice that we get a contribution from each current going through $R_2$. From the right loop, we get

$$V_2 - I_b R_2 + I_a R_2 - I_b R_3 - V_3 = 0$$

We now have two equations and two unknowns. So that we can check this in PSPICE, let’s put in some numbers. Set $V_1 = 25$ V, $V_2 = 9$ V, and $V_3 = 13$ V. Also, we’ll make $R_1 = 200 \Omega$, $R_2 = 60 \Omega$, and $R_3 = 170 \Omega$. Notice that your solution will give you 0.06121 A for
\(I_a\) and \(-0.0014235\, \text{A}\) for \(I_b\) (meaning \(I_b\) really goes counter-clockwise), and that these are the currents through batteries 1 and 3. The current through battery 2, therefore, is 
\[0.06121\, \text{A} - (-0.0014235\, \text{A})\], giving 0.06263 A. \(V_2\) is effectively being charged by the other two batteries.

**Supernodes and Supermeshes**

These methods (node & mesh) can be even easier in certain cases. For example, if there is a voltage source between two essential nodes, we can remove the source and consider the two nodes to be one supernode. This will serve the function of eliminating a variable (the current between the nodes making up the supernode) from the very beginning. Of course, at the end of the problem, you can always restore the voltage source to relate the potential between the nodes. You can think of this as being similar to the way you can solve the problem of a block on a frictionless table which is tied to another block hanging off the edge. You can consider each block to be its own isolated system (or node) with its own set of \(F = ma\) equations, or you can consider the two blocks & rope together to be a single system (supernode) and ignore the internal forces (voltages/currents) between them until the end.

Similarly, a supermesh is formed when a branch of a circuit containing a current source is mentally “removed” from the circuit. You can then write an equation that follows the meshes bordering the current source as if they were one. Of course, you still have two separate currents, but they appear in a single equation.

**Source Transformations**

We can also simplify circuits by changing the nature of some of the sources. A voltage source of \(V\) volts in series with a resistor of \(R\, \Omega\) can be replaced by a current source of \(I\) A in parallel with the same resistor. The relationship between \(V\) & \(I\) is just that found by Ohm’s law: \(V = I\, R\). To test this, look at the simple circuit below containing a voltage source, in series with a resistor, connected to some load resistor \(R_L\):

![Circuit Diagram](image)

The current through this resistor will be

\[I = \frac{V}{R + R_L}\]
If we replace the voltage source & resistor in series with a current source and resistor in parallel,

\[
\begin{align*}
\text{This leaves us with a current divider, meaning a current through } R_L \text{ of } \\
I_L &= \frac{IR}{R + R_L}
\end{align*}
\]

As you can see, this \( I_L \) will be equal to the current in the series circuit with a voltage source if \( V = IR \).

**Thevenin & Norton Equivalent Circuits**

We’ve already seen that we can take a collection of resistors in series and replace them with a single equivalent resistor, and we can also take a collection of resistors in parallel and replace them with a single equivalent resistor. This is true if we are only interested in the endpoints of these systems. Similarly, the Thevenin and Norton equivalent circuits extend those ideas so that, if we are interested in just two terminals of any complicated circuit made of resistors (series, parallel, both or neither), current sources, and voltage sources, we can replace that system with a single voltage source and single resistor in series with it (Thevenin equivalent). The Norton equivalent circuit is made by putting a single resistance in parallel with a single current source.

To make transformations to these equivalent circuits, we need two concepts: the open-circuit voltage at the terminals (i.e., what a perfect voltmeter would read if connected to the terminals) and the short-circuit current through the terminals (or what a perfect ammeter would read if connected to the terminals). We can get the Thevenin resistance from

\[
R_{Th} = \frac{V_{Th}}{I_{sc}}
\]

We’ve already seen how to turn a voltage source into a current source, so we can do that to get the Norton equivalent. We can also find the equivalent resistance by getting
rid of all the independent current and voltage sources. We replace voltage sources by wires (short circuit) and current sources by gaps (open circuits). Calculate the resistance seen at the terminals in this situation, and it will be the Thevenin resistance.

**Maximum Power Transfer**

If we have a Thevenin equivalent circuit with some \( V \) and \( R_{th} \), to what load resistance \( R_L \) can it most efficiently transfer its power? First, look at the case when \( R_L = 100 \times R_{th} \). \( R_L \) is large, so \( I^2R_L \) would be large, but \( I \) will actually be small because of the large series resistance. The small \( I \) is more important than the large \( R_L \) since it is squared. If instead we have \( R_L = 0.01 \times R_{th} \), the current will be approximately \( I_{sc} \) but the load resistance will be small. Decreasing \( R_L \) will have little effect on \( I \) at this point.

To see this mathematically, note that the power transferred to \( R_L \) will be \( I^2R_L \) where

\[
I = \frac{V_{th}}{R_L + R_{th}}
\]

To maximize the power transfer, we need to take the derivative of \( I^2R_L \) with respect to \( R_L \) and set it equal to zero. When you do that, you find \( R_L = R_{th} \). This general principle will be extended to the idea of *impedance matching* when we look at reactive circuit elements.

**Superposition Principle**

One of the most important properties of linear systems is the applicability of the superposition principle. You've seen this before in physics; if a mass is being acted on by three different forces, the net acceleration will just be the (vector) sum of those forces divided by the mass. They don't combine in any strange way; each force acts as if the others are not present.

We can use the same idea to analyze a circuit with multiple voltage/current sources. To do this, eliminate all current and voltage sources except one. Voltage sources are replaced by straight wires (short circuit) and current sources by gaps (open circuit). Analyze the new circuit (which has only one source) and record the values of currents through circuit elements and voltages across circuit elements. Don't forget to pay attention to the direction of currents and the polarity of the voltages. Now do the same for the next source (removing the first one and all others) and record currents & voltages. When this has been done for all sources, you can find the total currents and voltages by simple addition. For example, let's redo the problem below with the same numbers as last time. Set \( V_1 = 25 \text{ V} \), \( V_2 = 9 \text{ V} \), and \( V_3 = 13 \text{ V} \). Also, we'll make \( R_1 = 200 \Omega \), \( R_2 = 60 \Omega \), and \( R_3 = 170 \Omega \).
Going through this circuit, first get rid of $V_2$ and $V_3$, replacing them by straight wires:

We quickly see that $R_2$ and $R_3$ are in parallel and therefore can be replaced by an equivalent resistance of 44.35 $\Omega$, giving a total circuit resistance of 244.35 $\Omega$. The current through $R_1$ is 0.1023 A (to the right) and the voltage drop across it is $(0.1023 \text{ A} \times 200 \text{ $\Omega$}) = 20.46 \text{ V}$ (left end at higher potential). The voltage drop across $R_2$ is 25 V-20.46 V = 4.54 V (top at higher potential) giving a current of 0.07567 A (down) through it and $R_3$ has the same voltage drop of 4.54 V (left end at higher potential) causing a current of 0.02671 A (to the right).

Next, remove $V_1$ and put $V_2$ back in. We then find $R_1$ and $R_3$ in parallel with each other, the combination being in series with $R_2$. The total resistance = 151.89 $\Omega$, current through $R_2 = 0.05925$ A (up) with the potential drop being 3.56 V (bottom end at higher potential). This leaves a potential of 5.445 V across the $R_1$-$R_3$ pair with their connecting node being at higher potential than their outer ends. The current through $R_1$ is 0.02722 A (left) and the current through $R_3$ is 0.032 A (right).

Lastly, remove $V_2$ and reinstall $V_3$. $R_1$ and $R_2$ are now in parallel, giving a total circuit resistance of 216.2 $\Omega$ and a current through $R_3$ of 0.060 A (up). The potential across $R_3$ is 10.22 V (right end higher) meaning the potential across the other two resistors is 2.776 V (higher at connecting node). The current through $R_2$ is then 0.04626 A (down) and the current through $R_1$ is 0.01388 A (to the left).
What’s the final answer for the current through R₁? We choose the right to be positive, so we have 0.1023 A - 0.02722 A - 0.01388 A = 0.0612 A to the right (the same answer we found before, except for any rounding errors). The potential across this resistor will be 20.46 V - 5.445 V - 2.776 V = 12.24 V (left end higher). Notice that this satisfies V = IR for the resistor and that the current is flowing in the correct direction.

For R₂, we get a current (with up being positive) of -0.07567 A + 0.05925 A - 0.04626 A = -0.0627 A (meaning the current flows down) and a potential difference across it of 3.76 V with the top being at a higher potential. Finally, for R₃, the current is 0.02671 A + 0.032 A - 0.060 A = -0.00129 A giving a tiny current to the left. The potential difference between the ends (the right end must be higher or the current wouldn’t flow left) is 0.23 V.

This won’t simplify every circuit, but it will give current and voltage when the “problem” is just the presence of more than one source. The resistance of each component must be constant – not a function of voltage or current. Also, the components have to be bi-directional so that current flows the same way from either direction.

**Op Amps**

The ability to turn a small signal into a large signal (i.e., amplify it) has a variety of uses. In the past, this was done using vacuum tubes, and then transistors. Today, it is more common to use an operational amplifier, or op amp. The core of the op amp consists of two inputs and one output (there are additional connections to supply power for the amplification and, in some cases, connections to adjust tiny imperfections in the op amp’s behavior).

The inputs are known as the inverting input (represented by a minus sign) and the non-inverting input (represented by a plus sign). The op amp itself is drawn as a triangle (since the main features are two inputs and one output), as shown below:

![Op Amp Symbol](image)

The top left corner is the inverting input, the bottom left is the non-inverting input, and the corner on the right is the output. It’s common to leave out the external power connections when drawing the op amp symbol, but when included, they are usually placed as shown below:

![Op Amp Symbol with Power Connections](image)
For some op amps, the V- is replaced by a connection to ground. These are known as single-supply op amps, since these can be powered by a single source. The typical connection for a dual-supply op amp might be to connect two batteries together in series (+ of one to – of the other) and make that common connection the circuit’s ground. The free + terminal would then be V+ and the free negative terminal would be V-.

One of the important characteristics of an op amp is its gain (usually symbolized by A). This is the factor that multiplies the input voltages to arrive at the output voltage. This number can be greater (sometimes much greater) than a million. The actual gain can only be achieved when the op amp is operating in a small range of voltages; we can’t take an op amp powered by a 9 V battery, put a 3 V signal on the + input, and expect a voltage of 3,000,000 V at the output. A single-supply op amp powered by a 9 V battery can’t output a signal lower than ground or higher than 9 V above ground. Depending on the particular model, the op amp may only be able to get within a volt or two of the power supply’s voltage. In other words, we might see that an op amp powered by the 9 V battery above never outputs a signal lower than (for example) 1.5 V above ground and never higher than 7.5 V above ground. Because the two terminals of the battery may be connected to long conductors (known as rails) to give us multiple points of connection to either 9 V or ground, an op amp that can get very close to the limits (ground and 9 V here) is known as a rail to rail op amp.

The output of the op amp is a function of the gain A and the voltages on the two inputs V+ and V-. Ideally, it will be A (V+ - V-). If the voltage difference between the inputs (V+ - V-) is large enough for the preceding formula to predict a voltage larger than the op amp’s positive voltage limit or smaller than its negative voltage limit, the op amp is saturated and the output will be capped at the appropriate limit. If the formula does not predict saturation, the op amp is said to be operating in the linear region.

Because the gain for an op amp is typically so large, it will only be in the linear region for input voltage differences smaller than (Vsupply+ - Vsupply-)/A. In other words, V+ and V- will have to differ by much less than a volt (typically mV to µV). Given the precision of most voltmeters, if we assume the op amp is in the linear region, we can usually make the ideal op amp approximation that the voltages on both inputs are equal:

\[ V_+ = V_- \quad \text{or} \quad V_p = V_n \]
Another part of the ideal op amp approximation is that the input impedance is infinitely large. This means that the inputs will not draw any current. The current leaving the output will come from the power supply ($V_{\text{supply}}^+$ or $V_{\text{supply}}^-$).

**The Inverting Amplifier**

In the figure below, an op amp is used to make an **inverting amplifier**.

Because the incoming voltage $V_S$ is connected (through resistor $R_S$) to the inverting input, the output will be negative with respect to ground. Using the ideal op amp rule that the current into the inverting input is zero, and the voltage of the inverting and non-inverting outputs is the same (and therefore zero, since the non-inverting input is grounded), we can use the node voltage method to investigate the currents leaving the inverting input. We get

$$\frac{V_{\text{inv}} - V_S}{R_S} + \frac{V_{\text{inv}} - V_O}{R_f} = 0$$

where $V_O$ is the output voltage and $R_f$ is the resistance connecting the output to the inverting input. Using $V_{\text{inv}} = 0$ as mentioned before, we get

$$\frac{-V_S}{R_S} = \frac{V_O}{R_f} \Rightarrow V_O = -\frac{R_f}{R_S}V_S$$

where $R_f$ may be much larger than $R_S$. Of course, if $R_f/R_S$ is larger than the gain $A$ of the amplifier, this equation can’t hold. For an ideal amplifier, $A$ is infinite, so it’s not a problem. The resistor $R_f$ is a source of **negative feedback** to the op amp. This is very important for stability. If the output happened to fluctuate upward just a little, that increased voltage would then appear across the $R_f$-$R_S$ voltage divider. Because of the
difference between $R_f$ and $R_s$, this would amount to a tiny voltage change at the inverting input. That voltage change gets magnified and inverted, so that will bring the output back down.

Negative feedback acts as a restoring force on the op amp’s output, sort of like the way a ball inside a bowl feels a restoring force from the combination of gravity and the shape of the bowl; if the ball moves from the center to the right, a force pushes it back towards the center. In both cases, this restoring force lends stability to the situation. Without negative feedback, the (typically) huge gain of the op amp will tend to drive it to its limit (near $V_+$ or $V_-$, depending on the voltage difference between the inverting and non-inverting inputs).

**Summing Amplifier**

The same ideas can be applied to an amplifier which has multiple inputs to the inverting input. In the figure below, we have three signals $V_a$, $V_b$, and $V_c$ each connected to the inverting input through resistors $R_a$, $R_b$, and $R_c$. If we again use the ideal op amp assumption, we see that $V_{\text{inv}}$ has to be the same as $V_{\text{non-inv}}$, which is at 0 V, and the current coming into the inverting input again has to be zero. Through the node-voltage method, we now get

$$\frac{V_{\text{inv}} - V_a}{R_a} + \frac{V_{\text{inv}} - V_b}{R_b} + \frac{V_{\text{inv}} - V_c}{R_c} + \frac{V_{\text{inv}} - V_O}{R_f} = 0$$

The substitution $V_{\text{inv}} = 0$ then gives

$$V_O = \left(\frac{-V_a}{R_a} + \frac{-V_b}{R_b} + \frac{-V_c}{R_c}\right)R_f$$

The output is a scaled version of the (inverted) weighted sum of the input voltages. If $R_a = R_b = R_c$, the input voltages are weighted equally.
Noninverting Amplifier

Another amplifier design involves sending the input signal $V_g$ through a resistor $R_g$ into the noninverting input. From the circuit below, if $V_o$ goes through $R_f$ and $R_S$ to ground, the inverting input sees the middle of a voltage divider circuit. That voltage is then

$$V_{inv} = \frac{V_o}{R_f + R_S} R_S$$

If the inverting and noninverting inputs are at the same voltage, $V_g = V_{inv}$, so we get

$$V_O = \frac{R_S + R_f}{R_S} V_g = V_g \left(1 + \frac{R_f}{R_S}\right)$$

as long as the op amp is in the linear gain region (i.e., $A < 1 + R_f/R_S$).
The difference amplifier

The circuit below combines elements of the inverting and noninverting amplifiers to form a difference amplifier.

The noninverting input features a voltage divider configuration which gives a $V_{\text{Non-inv}}$ of

$$V_{\text{Non-inv}} = \frac{R_d}{R_c + R_d} V_b = V_{\text{inv}}$$

where the second equality uses the ideal op amp assumption. The node voltage method also tells us that the voltage on the inverting input should satisfy
\[ \frac{V_{\text{inv}} - V_a}{R_a} + \frac{V_{\text{inv}} - V_O}{R_b} = 0 \]

Combining these two equations yields

\[ V_O = \frac{R_d (R_a + R_b)}{R_a (R_c + R_d)} V_b - \frac{R_b}{R_a} V_a \]

If the resistors are chosen so that

\[ \frac{R_a}{R_b} = \frac{R_c}{R_d} \]

then we get that

\[ V_O = \frac{R_b}{R_a} (V_b - V_a) \]

therefore amplifying the difference between the input voltages.

The two input voltages can also be represented as a linear combination of \( V_a \) and \( V_b \), specifically

\[ V_{\text{dm}} = V_b - V_a \quad \text{and} \quad V_{\text{cm}} = \left( V_a + V_b \right) / 2 \]

where \( V_{\text{dm}} \) stands for the differential mode input and \( V_{\text{cm}} \) is the common mode input. By rewriting these two equations to solve for \( V_a \) and \( V_b \) in terms of the common mode and differential mode voltages, we can return to the equation for the output voltage of a differential amplifier and write it as

\[ V_O = \left[ \frac{R_a R_d - R_b R_c}{R_a (R_c + R_d)} \right] V_{\text{cm}} + \left[ \frac{R_d (R_a + R_b) + R_b (R_c + R_d)}{2 R_a (R_c + R_d)} \right] V_{\text{dm}} \]

The terms in square brackets can also be written as \( A_{\text{cm}} \) and \( A_{\text{dm}} \), the common mode gain and the differential mode gain. For an ideal difference amplifier, \( A_{\text{cm}} = 0 \) and even small differences between the inputs are amplified. One of the measures of op amp performance is the ability of the op amp to amplify the difference between input voltages.
and ignore the average of them (the part common to both inputs). This is called the **common mode rejection ratio**, or CMRR, and it’s calculated as

\[
CMRR = \left| \frac{A_{dm}}{A_{cm}} \right|
\]

where a larger value is more desirable.

**Real Op Amps**

When working with real-world op amps, things are a little different than our ideal. The inputs now have finite resistances rather than infinite. A voltage into these inputs now means the input currents are not zero, so we can’t use \( i_{\text{inv}} = i_{\text{non-inv}} = 0 \) anymore. Also, there is a finite output resistance (rather than zero), so there is a limit to the load they can drive. A more realistic model of the op amp (from Fig. 5.15 in your book) is shown below (with the ± V connections to the power supply omitted for now):

![Real Op Amp Model](image)

In this model, we can now see the \( R_i \) and \( R_o \) resistances, as well as the dependent voltage source (it depends on the voltages on the inputs). Your book re-analyzes the inverting amplifier under these new assumptions, and the node voltage method is applied to the inverting and non-inverting inputs to get

\[
\begin{align*}
\frac{V_{\text{inv}} - V_S}{R_S} + \frac{V_{\text{inv}}}{R_i} + \frac{V_{\text{inv}} - V_O}{R_f} &= 0 & \text{and} \\
\frac{V_O - V_{\text{inv}}}{R_f} + \frac{V_O - A(-V_{\text{inv}})}{R_O} &= 0
\end{align*}
\]
where, as in the original discussion of the inverting amplifier, \( R_f \) is the resistor connecting the output to the inverting input and \( R_s \) connects the input signal \( V_s \) to the inverting input. A little math reveals that

\[
V_O = \left( \frac{-A + \left( R_O / R_f \right)}{R_s \left( 1 + A + \frac{R_O}{R_i} \right) + \left( \frac{R_s}{R_i} + 1 \right) \frac{R_O}{R_i}} \right) V_s
\]

For comparison with the ideal case, your book lists values of \( R_o = 75 \ \Omega \), \( R_i = 2 \ \text{M} \Omega \), and \( A = 100000 \) for the 741 op amp (a very common model). A similar analysis of the non-inverting amplifier is also in your book.

**Bibliography**

*Electric Circuits*, 7th edition, Nilsson & Riedel
*Basic Circuit Analysis*, 2nd edition, John O’Malley
*Electric Circuits*, 4th edition, Nahvi & Edminister
http://www.allaboutcircuits.com - Tony Kuphaldt
http://focus.ti.com/lit/an/slod006b/slod006b.pdf - Ron Mancini
http://inst.eecs.berkeley.edu/~ee40/fa06/lectures/cch-LecturesF06.pdf - Connie Chang-Hasnain