AN INTRODUCTION TO ELECTRICAL ENGINEERING

Douglas De Boer, 2018

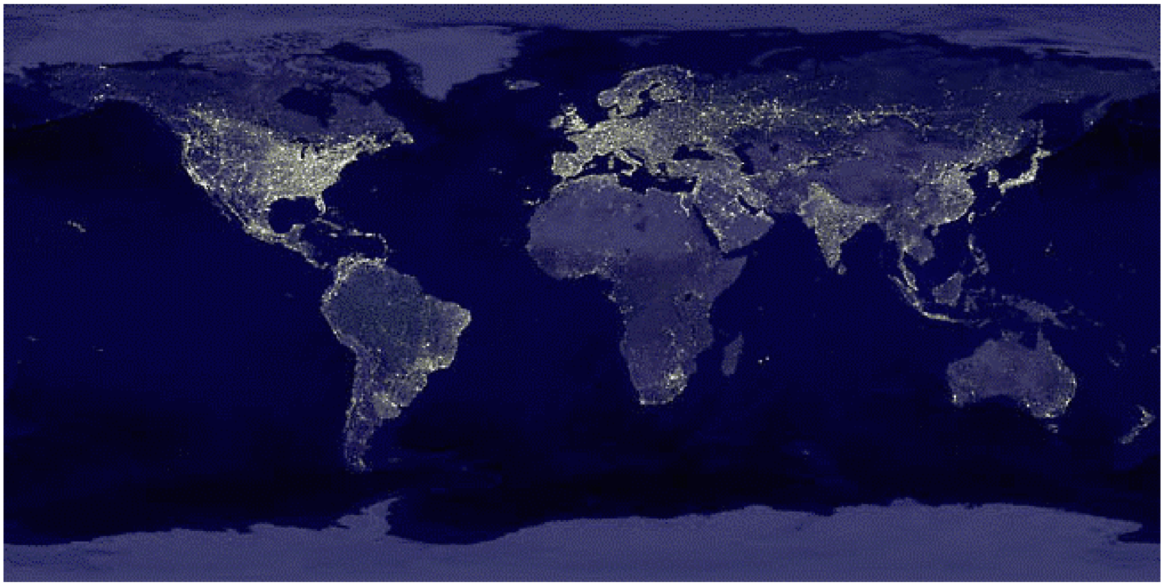


Figure 1. The Earth at night, a photomontage. [1]

Imagine that you come from another galaxy and are approaching the solar system on an exploratory mission. As you inspect the planets of the Solar System by means of your telescope you find that the third planet from the sun exhibits spectacular bright areas on its dark side. Further investigation shows that many electromagnetic signals bearing interesting codes also emanate from this planet. This, the third planet from the sun, seems especially worth exploring! [2]

As Figure 1 makes obvious, electricity plays an important role in our world's cultures. In the future, we will probably depend even more on electricity. Why is this? And just what is electricity?

1. BASICS: CHARGE AND CURRENT

Electricity has been studied since prehistoric times, although it was not recognized as "electricity" until about the 1600's. In antiquity the behavior of magnetized pieces of ore, called "lodestones," and static electricity caused by friction were great curiosities. These phenomena remain the typical introduction to electricity as written in modern physics books. The key developments leading to the use of electricity as we know it happened after 1600.

Early investigations of electricity involved the generation of static electricity by friction. Eventually the early theories of an "electric fluid" were supplanted by the modern theory of "charged particles" used to explain the forces of attraction and repulsion that are observed.

Definition: Electric ***charge*** is a fundamental property of matter, specifically of sub-atomic particles. It exists in two kinds, positive and negative. Similarly charged particles repel each other and oppositely charged particles attract each other. Engineers symbolize an amount of charge with the letter Q or q or $q(t)$.

Definition: ***Electromagnetism*** is the theoretical framework used to explain the physical phenomena arising from the motion and accumulation of electrically charged particles.

The forces related to the behavior of electrically charged particles are called electromagnetic forces. Electromagnetic force is one of four fundamental forces in the universe. The other three fundamental forces are gravity, the strong nuclear force (holds atoms together), and the weak nuclear force (causes radioactive decay).

Returning to the question of "Just what is electricity?" we find that there is no single precise definition. In 1820 Hans Oersted discovered that electric current in a wire produces a magnetic field. A year later Michael Faraday discovered that a moving magnet could produce a current in a wire. From then on electricity and magnetism have been understood to be linked together. In 1864, James Clark Maxwell developed a set of four equations which explain the observed linkages between electric current and magnetism. The modern perspective of physicists is that what we call electricity and magnetism are merely some of the behaviors of the same underlying electromagnetic phenomena. Other electromagnetic phenomena are radio waves, heat, light, and x-rays. It might be better to think of electrical engineers as "engineers of electromagnetism." Indeed, electrical engineers design lighting techniques, computer disk drives (both magnetic and optical types), fiber optic systems, medical and industrial x-ray equipment, heat sinks, temperature sensors, refrigeration systems (possibly based on the Peltier effect—direct conversion of electrical energy into a temperature differential with no compressor or other mechanical apparatus), photovoltaic solar collectors, and much more. The field of electrical engineering spans the gamut of topics in electromagnetism.

Benjamin Franklin is credited with naming the two kinds of charge "positive" and "negative." Franklin thought of "positive" as an excess of some kind of electric fluid and "negative" as a deficiency of the fluid. The idea of an electric fluid is now obsolete. The modern view is that any bulk of matter in its neutral state contains equal amounts of two kinds of charged particles, positive and negative. In that case we say that the algebraic sum of charge is zero. If two bodies of different material are rubbed together, then some charges will move from one body to the other body. This will upset the balance of charge on each separate body but leave the net algebraic sum of charge on the two bodies still at zero. This theory has been found satisfactory to explain all the familiar phenomena of static electricity. That the net balance of charge on the two bodies remains at zero (or balanced) expresses the concept that charge is conserved. Charged particles cannot be created or destroyed, they can only be moved from place to place.

Besides friction, there are also other ways to make charge move, not necessarily upsetting the balance between positive and negative charge in a particular body. For example, it is possible to swirl the charge around inside a stationary object. A changing magnetic field, as might be caused by a magnet in motion, is a very important way to cause charge to move in metals without upsetting the balance of charge within the metal.

We use the symbol Q for a constant amount of charge and the symbol q or $q(t)$ for a charge that varies with time. In general, we will use upper case variables to represent constant or average amounts of a quantity and lowercase variables to represent the instantaneous amount of a quantity at a particular time when the amount varies as time goes by.

The unit of measure for charge is the *coulomb*, abbreviated as "C." For example one might write, "after scuffing my feet on the carpet a measurement revealed that my body had a $Q = 0.01$ C. You could read that equation as "charge equal to one one-hundredth of a coulomb." The unit for measuring electric charge, the coulomb, is actually defined in terms of electric current, so we will consider the definition of current next, before considering the formal definition of the coulomb.

There are always several attributes of a thing that we can measure, and we can usually debate how we will measure these attributes. For example, how big is a house? Shall we say 1800 square feet, two stories, or eight rooms? As another example, consider the measurement of a current in a river. A boater probably would want to know the speed of the surface of the water with respect to the bed of the river in order to understand better how to pilot the boat to a destination up or downstream. Navigating in a 1 mile-per-hour current would be much different than navigating in a 20 mile-per-hour current! The speed of the surface of the water might vary along the path of intended travel of the boat. Factors such as the depth, width and curve of the river and the presence or absence of obstructions such as boulders in the river could all change the speed of the surface of the water. On the other hand, the manager of a reservoir who wants to know how fast the reservoir is being filled by the river would want to know the volume flow rate in some units like liters per second. That measurement would be the same at all points up and down the river regardless of the river's depth, width, curve, or obstructions as long as there are no other tributaries flowing into or out of the river. Both types of measurement somehow seem to quantify the amount of current in the river, but in different ways and for different purposes. Similarly, one could conceive different ways to measure electric current. In order for measurements of electric current to be understood and useful, we must agree on a standard definition that matches the intended use of the measurement. The chosen definition for electric current is analogous to the type of definition one would use to understand how fast a reservoir could be filled. The definition is not based on the speed of the charged particles but instead is based on the quantity of charge per second that is being moved.

Definition: Electric ***current*** is the quantity of net positive electric charge per unit time that flows in a given direction past a reference point. Symbol I or i or $i(t)$. Mathematically, $i = dq/dt$ or using constant amounts, $I = Q/T$, where t or T represent time and d represents the differential operator.

In the definition above, the phrase "given direction" is usually referring to an arrow on a circuit diagram. The arrow labels the "given direction." The phrase "net positive electric charge" means that the definition applies when the charge carriers are positive charges moving in the given direction. If positive charges are actually moving in the opposite direction to the given direction, then the current is a negative amount. If the charge carriers are actually negative charges, then the direction of "net positive" current flow is considered to be opposite to the direction of motion of the negatively charged carriers. For example, if electrons, which are negative charges, are flowing to the left (net positive charge flow is to the right), and the given direction is to the right, then the amount of current will have a positive sign. If electrons are flowing to the right ("net positive charge flow is to the left) and the given direction is to the right, then the amount of current flow has a negative sign.

The equation $i = dq/dt$ can be understood as "current equals a small amount of net positive charge (in coulombs) flowing past a point in a given direction divided by the small amount of time it took to flow (in seconds)." The symbol "d" (as in dq and dt) is used to represent the concept of making the measurement as a ratio of small quantities. In other courses you may see the symbol Δ (delta) used for this idea of a ratio of small quantities, or small differences, as in $i = \Delta q/\Delta t$. The difference between Δ and d is that d implies that the ratio gets more precise as the quantities (in particular the denominator) approach but never quite get to zero. This is especially true with respect to knowing precisely the time at which the amount of current was flowing. We are assuming here that the measurements of the small quantities dq and dt are infinitely accurate in the first place so that measurement errors do not create problems.

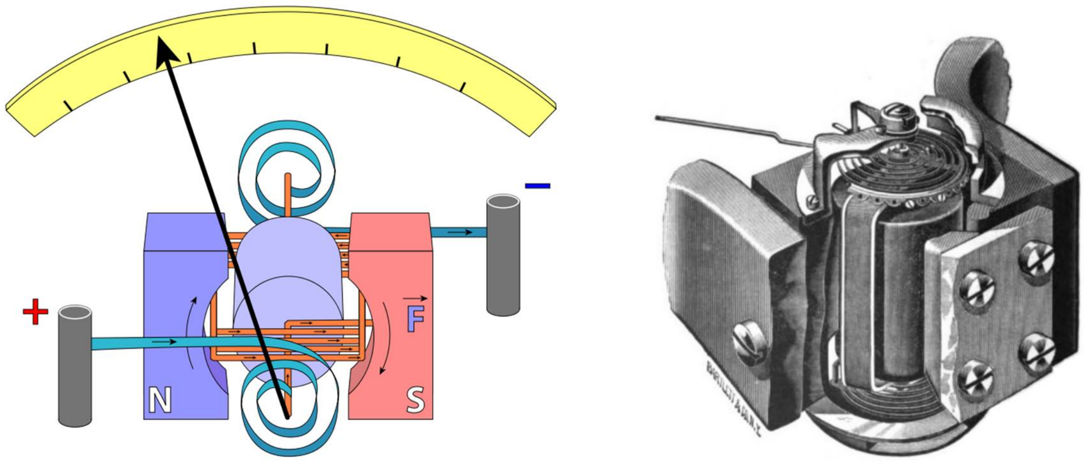
As a practical matter, measuring small quantities such as dq and dt cannot always be done with infinite accuracy. Sometimes relatively large quantities have to be used in the ratio Q/T. Then we get the average of the current over the time interval represented by T, where T represents the time the charge flow Q started subtracted from the time when the measurement stopped. ($T = t_{\text{stop}} - t_{\text{start}}$) In that case we can write $I = Q/T$. The equation can be algebraically manipulated and it remains true. Thus $Q = IT$ and $T = Q/I$.

The letter C is not used to symbolize current or charge since it is used by convention for a quantity called capacitance. Therefore the symbol for a quantity of charge is Q or q or $q(t)$ and for current I or i or $i(t)$. The abbreviation for coulombs, the SI unit for measuring charge, is also the letter "C." We distinguish between abbreviations for units and symbols for variable quantities by using italics for symbols but not for abbreviations of units. For example, 5 C is five coulombs, but $5C$ is five times the unknown amount of capacitance represented by the symbol (or variable) C.

Consider now the unit for measuring current, the ampere. The work of Hans Oersted and James Clark Maxwell showed that current produces a magnetic field. Thus two current carrying wires suspended near each other will either be attracted or repulsed from each other depending on the direction of current flow. This force is the basis for the definition of the ampere.

Definition: The ***ampere*** is the unit of measure for electric current in the International System (SI) of units. One ampere is the constant current that if maintained in two parallel straight conductors (of infinite length and negligible circular cross section) separated by a distance of one meter in a vacuum, produces a force between the wires of 2×10^{-7} newtons per meter of length. Also known informally as an amp. Abbreviated as "A."

The important point to remember in the above definition is that current is measured by means of a related force. In a practical analog current meter, the current flows through a small coil located near a permanent magnet. The resulting force is used to turn a pointer against a spring's torque. There are several geometries used for these meters. A very popular geometry is called a d'Arsonval meter movement. These analog mechanisms are still in use in many places. Being directly related to the definition of the ampere, these analog current meters can be surprisingly precise. Accuracy of better than 5% is a cinch. Accuracy of 2% or 1% is typical of a moderately priced instrument. Much better is possible via elaborate mechanisms. Of course there are electronic digital meters that also can measure current, but to calibrate them, one must ultimately refer back to a meter which operates on principles related to the definition of the ampere—an analog meter.



(a) (b)

Figure 2. Mechanism of a d'Arsonval meter movement. (a) Schematic, (b) Cut-away illustration.[3]

2. SI UNITS FOR ELECTRICAL QUANTITIES

The ampere is one of seven *basic units* in the SI system. A basic unit is a unit whose definition relies on some arbitrarily chosen thing. As an example, at one time there was a "meter-stick" kept in France which was considered by international agreement to be the definition of the meter for the whole world. (You had to use it at a temperature of exactly 0 C because of thermal expansion!) If you wanted to test your tape measure or other distance measuring device and have the most accurate test possible, you would have to go to France and compare your measuring device to the standard meter-stick. That meter-stick is no longer in use. The meter is now defined in terms of the speed of light, since subsequent to Einstein's theory of relativity, we assume the speed of light is a constant of nature. The definition was chosen to be as close as possible to the length of the original standard meter-stick. The unit of mass is still defined in terms of a particular mass of platinum-iridium kept in France. The National Institute of Standards and Technology (NIST) keeps a very precise copy for use in the United States. All seven basic SI units refer back to some chosen artifact which is assumed to be unchanging. The seven arbitrarily defined basic units in the SI system are:

Basic Unit	Abbreviation	Base Quantity
meter	m	length (or distance)
kilogram	kg	mass
second	s	time
ampere	A	electric current
kelvin	K	temperature
candela	cd	luminosity
mole	mol	amount of a substance

Table 1. The seven basic SI units.[4]

Although the definition of the ampere requires the prior definition of the meter and the newton, the arbitrary selection of the distance between conductors, the amount of force, etc. is what makes the definition basic. All SI units not on the list in Table 1 are said to be *derived* units. Derived units are always ratios and/or products of the seven basic units (or other derived units).

The ampere is the only basic SI unit that relates directly to electricity or electromagnetism. All the other electrical and magnetic units in the SI system are derived from the seven basic SI units. For example, the definition of the coulomb is derived from the definitions of the ampere and the second.

Definition: The ***coulomb*** is the SI unit of measure for electric charge. One coulomb is the amount of net positive charge that flows past a given point in a given direction in one second if the flow of current is one ampere. Abbreviated as "C."

The definition above shows that the number of coulombs can be found by multiplying the number of amperes times the number of seconds during which the current flows. (coulombs are "amp-seconds"). It is a product of two basic SI units. This makes the coulomb a derived SI unit.

A charge of –1 C happens to correspond to about 6.241506×10^{18} electrons. Although that is not the definition of a coulomb, it is about right to say that a coulomb is a charge equal to the magnitude of the charge of 6.241506×10^{18} electrons. (Keep in mind that electrons are negatively charged.)

3. VOLTAGE AND VOLTS

It takes work to make charged particles move. This is because like charges repel each other. Making the charge move against a force of repulsion takes work.

Definition: ***Work*** is a force applied through a distance. Symbol W or w or $w(t)$. Mathematically, if the force is constant, $W = FX$ where F is a constant force in the direction of the distance and X is the distance.

Definition: A ***newton*** is the SI unit of measure for force. A force of one newton is that force which causes a mass of one kilogram to accelerate at a rate of one meter per second squared. Abbreviated as "N."

Definition: The ***joule*** is the SI unit of measure for work (and energy). One joule of work is done by a constant force of one newton acting over a distance of one meter. Abbreviated as "J."

The amount of work needed per unit charge to make the charge move is called *voltage* or *electric potential*, or *electromotive force*. *Voltage* is now the commonly used term. Voltage is measured in *volts*. This is a bit of an unfortunate choice of language. We do not speak of footage. Instead, we speak of length or distance and might measure that in feet. It helps to remember that there are two concepts embodied in similar words—the phenomenon of *voltage* having to do with subjects like work, distance and force and the unit of measure, the *volt*.

Definition: ***Voltage*** is the amount of work needed per unit of charge to make charge move from a reference point (labeled "–") to another point (labeled "+"). Symbol V or v or $v(t)$. Mathematically, $v = dw/dq$ or using constant amounts, $V = W/Q$.

Definition: The ***volt*** is the SI unit of measure for voltage. One volt is one joule per coulomb. Abbreviated as "V."

Notice that since voltage implies a distance, it is always measured with respect to *two* points. For example, we can measure the voltage *across* the two terminals of a battery. We can also measure the current that flows through the battery. It is helpful to associate the word "across" with the words "voltage" and "volt." (It is *incorrect* to say that you "measured the voltage *through* the battery.") Voltage is *across* a pair of points in contrast to current, which flows *through* something.

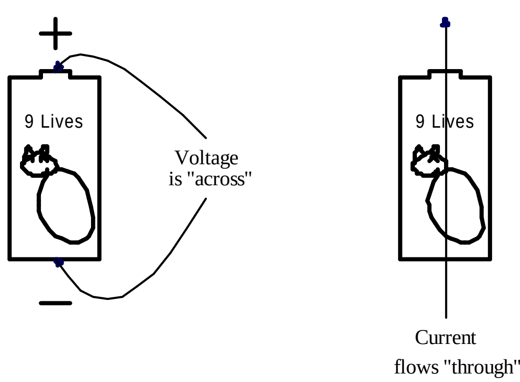


Figure 3. Voltage is "across" and current flows "through."

Example set #1

On a cold winter day the average voltage of a certain car battery is 12 V. The car's starter motor must operate for five seconds to start the car. The starter draws 100 A. How much charge flows through the starter to start the car?

From the definition of current, $I = Q/T$. Solving for the charge, Q, gives:

$$Q = IT$$
$$Q = (100 \text{ A})(5 \text{ s}) = 500 \text{ C}$$

How much work does the starter do to start the car?

The word "work" in the question reminds us of voltage, the amount of work done per unit charge. From the definition of voltage, $V = W/Q$. Solving for the work, W, gives:

$$W = QV$$
$$W = (500 \text{ C})(12 \text{ V}) = 6000 \text{ J}$$

Suppose that this car was to be started by hand with a crank. The handle of the crank is at a radius of $R = 1/\pi$ m (about one foot) from the center of rotation. The crank is turned at a speed of two revolutions per second and that the car starts at the end of the tenth revolution of the crank, that is, after 5 seconds. How much force must be applied to the handle of the crank to start the car?

We know it takes 6000 J of electrical work to start the car. We will assume the starter motor is 75 percent efficient so that the mechanical work that needs to be done to start the car is only 4500 J. Set that equal to the amount of work the crank must do, which is the force applied to the handle, F, times the distance the handle moves, X. In five seconds the crank will go around the circumference of a circle ten times and the circumference of a circle is 2π times the radius, so

$$X = 10(2\pi R) = 10[2\pi(1/\pi \text{ m})] = 20 \text{ m}.$$

Using the definition of work and solving for F gives:

$$W = FX$$
$$F = W/X = (4500 \text{ J})/20 \text{ m} = 225 \text{ N}$$

Converting that to pounds-force gives about 50 lbf (The numbers in this example are about right for a big car or SUV.)

4. ELECTRIC ENERGY IS EASILY TRANSMISSIBLE BUT DIFFICULT TO STORE

Electric power has one distinct advantage over most other types of power or energy. It can be easily transported. Truly large amounts of power can be transported by wires over hundreds and even thousands of miles. Small amounts of power that bear information can be sent by radio waves over practically any distance, even an intergalactic distance. Electricity also has a disadvantage: It is difficult to store.

Power is the rate at which work is done. Electric power requires both voltage and current. The amount of electric power flowing at any instant of time can be found by multiplying instantaneous voltage times the instantaneous current.

Definition: ***Power*** is the rate at which work is done. Symbol P or p or $p(t)$. Mathematically, $p = dw/dt$ or using constant amounts $P = W/T$

Definition: The ***watt*** is the SI unit for power. One watt is one joule per second. Abbreviated as "W."

Theorem: At any instant in time $\boldsymbol{p = vi}$.

The theorem above can be proved. Start with the definition of voltage, $v = dw/dq$, and the definition of current $i = dq/dt$. The product is:

$$vi = (dw/dq)(dq/dt) = dw/dt$$
but $dw/dt = p$ by the definition of power.
Thus, $p = vi$

Example Set #2

Suppose a 30 mile long power transmission line operates at 5000 V and caries a current of 100 A. Also suppose that a typical farm uses an average amount of power of 10000 W. How many farms can be powered by this transmission line?

The power that can be delivered by the transmission line is

$$P = VI = (5000\text{ V})(100\text{ A}) = 500000\text{ W}$$

The number of farms that can be powered can be found by dividing by the average load of a farm

Number of farms = 500000 W/10000 W per farm = 50 farms

The zeros in the numbers above make them a little hard to read. Below the problem is restated using engineering notation (k = kilo, M = mega, etc.). Electrical engineers frequently use engineering notation to keep numbers easily readable.

$$P = VI = (5\text{ kV})(100\text{ A}) = 500\text{ kW}$$

Number of farms = 500 kW/10 kW per farm = 50 farms

Multiple	Prefix	Abbreviation	Multiple	Prefix	Abbreviation
10^{24}	yotta	Y	10^{-3}	milli	m
10^{21}	zetta	Z	10^{-6}	micro	μ
10^{18}	exa	E	10^{-9}	nano	n
10^{15}	peta	P	10^{-12}	pico	p
10^{12}	tera	T	10^{-15}	femto	f
10^{9}	giga	G	10^{-18}	atto	a
10^{6}	mega	M	10^{-21}	zepto	z
10^{3}	kilo	k	10^{-24}	yocto	y

Table 2. Engineering prefixes used with SI units.

5. "FREE NIAGARA"

Starting in about 1890 there were growing demands to use the power of Niagara Falls for industrial purposes. After about 75 years of development ending in the mid 1960's the potential power generation capabilities of the Falls had been fully realized. The story of the development of hydro power at Niagara Falls is a story of protecting the environment from a mess even uglier than the dams, reservoirs, and hydroelectric generating stations we know today.

Back in the 1800's each factory or mill that derived power from the Falls had its own water-wheel or turbine. Canals were dug from the upper part of the river to the factories to supply water for the water-wheels and turbines. Due to the landscape the canals ran over, only a fraction of the available head of water could be used for power. Each factory transmitted the power from its water wheel or turbine into the factory on a steel shaft. This main shaft drove other secondary shafts by means of gears, belts and pulleys that were typically mounted along the ceiling of the factory. Ultimately, every machine in the factory attached to this system of shafts by means of a belt. (See Figure 4 on the next page.)

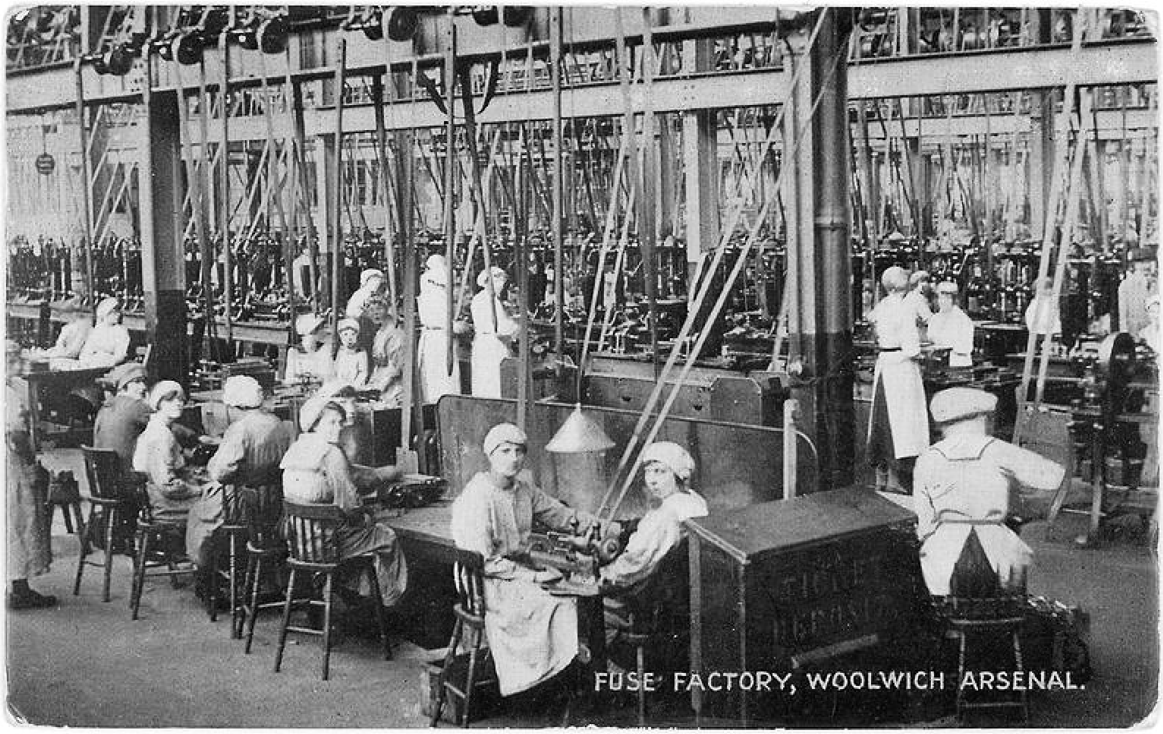


Figure 4. A typical factory scene in the 1800's. Note the overhead shafts. Power is delivered to each machine from an overhead shaft via a belt.[5]

Figure 5. The mill district on the lower Niagara River in the 1880's.[6]

Figure 6. ***A Distant View of Niagara Falls***
An artist's conception of what a pristine Niagara Falls might look like.[7]

There were many canals and factories near the Niagara Falls waterfront. The people living in Buffalo New York and surrounding areas wanted to do something about this industrial blight on the scenic landscape of the Falls. Moreover, growth was stymied since it was not practical to build canals to deliver water to distant factories. (See Figure 5.) A cry went up to "free Niagara." European and American companies competed to design a way to transmit the power from the river to factories located a mile or more from the river. Of the final proposals, four were for pneumatic systems, two for hydraulic systems, and one for wire rope running over pulleys. Pneumatic systems are notoriously inefficient since air is compressible, but they are relatively safe. Hydraulic systems are more efficient, but also more dangerous since a spray of hydraulic oil from a leaking pipe is flammable. The idea of a wire rope system would have been something like a super high-speed ski lift, without the chairs, where the pulleys on the towers could be attached to the factory's shafts to tap power. In retrospect, this just seems like a silly way to transmit power in a city environment, not to mention dangerous! A broken wire rope could fall to the ground between the towers and whip along the sidewalk, mowing down every pedestrian in its path. Eventually, there were also seven proposals for electric systems.[2] Electric systems were chosen since only they could scale up to transmit megawatts or even gigawatts of power with relative safety and efficiency.

Today, counting power generated on both the American and Canadian sides of the river, there is about 4.4 GW of hydroelectric power available from Niagara Falls. That is enough to light 44 million 100 W light bulbs. Equivalently, that power replaces about 5 or 10 coal-fired electric generating stations of today's typical sizes. Imagine trying to transmit that power and distribute it in New York City and Toronto via shafts and gears spanning the miles, or compressed air, or wire rope and pulleys! It is impossible to conceive. Electricity uniquely makes it possible. The only present-day competition to electricity as a means of transmitting power from one place to another is provided by oil and natural gas via pipelines and tanker ships or via coal in train cars.

Figure 7. Lewiston reservoir (top), Robert Moses Power Station (middle) and Sir Adam Bcck Power Station No. 2 (bottom left) along the lower part of the Niagara River. [8]

The main disadvantage to electrical power has to do with the difficulty of storing it as electricity. Only two devices do this, inductors and capacitors. Batteries store chemical potential energy in a way that can be readily converted to and from electrical energy, but even they cannot compete with the energy density of oil and gasoline, or even coal. This is true both in terms of energy stored per unit mass and energy stored per unit volume. These are important restrictions. Battery operated electric cars are a present-day possibility, but a battery operated electric commercial airliner is out of the question.

The problem of energy storage is also an issue at Niagara Falls. The electricity generated at the Falls is needed mostly in the daytime, especially in the afternoons and early evenings. Diverting more of the river from the falls and into the turbines in order to generate more power in the afternoons would diminish the beauty of the falls at precisely the times when the most tourists are present. Of course in the wee hours of the mornings few people are looking at the falls, so it would be nice to generate power then and store it for the next day's use. That is not directly possible however. Instead, reservoirs such as the Lewiston Reservoir (Figure 7), are filled at night. Then they are emptied during the daytime to augment the flow of river water into the turbines to generate more electricity while also allowing the falls to flow at a normal volume during the times when tourists are present. (There are reservoirs on the Canadian side of the river too.)

The present-day tourist experience at Niagara Falls is indeed a highly engineered experience. Some people (technological positivists) view technology as the answer to all our problems. Indeed, hydroelectric power has allowed the industrial blight along the Niagara River to be cleaned up. On the other hand, the reservoirs have their own risks and environmental consequences. The history of the technological development of Niagara Falls is a quintessential case study showing how tightly intertwined technology is with cultural development. Psalm 8 tells us that we have been made "ruler over the works of [Jehovah's] hands." what responsibility!

1O LORD our Lord, how excellent is thy name in all the earth!
who hast set thy glory above the heavens.

2Out of the mouth of babes and sucklings hast thou ordained strength
because of thine enemies, that thou mightest still the enemy and the avenger.

3When I consider thy heavens, the work of thy fingers,
the moon and the stars, which thou hast ordained;

4What is man, that thou art mindful of him?
and the son of man, that thou visitest him?

5For thou hast made him a little lower than the angels,
and hast crowned him with glory and honour.

6Thou madest him to have dominion over the works of thy hands;
thou hast put all things under his feet: 7All sheep and oxen,
yea, and the beasts of the field; 8The fowl of the air, and the fish of the sea,
and whatsoever passeth through the paths of the seas.

9O LORD our Lord, how excellent is thy name in all the earth! (Psalm 8, KJV)

In the future, fossil fuels (coal, oil, natural gas) will continue to become more scarce and valuable. They will increasingly be reserved for purposes that exploit their value best, such as fuel for jet airplanes, and the manufacture of petrochemicals, pharmaceuticals, and plastics. In the future we will not be able to afford to burn valuable oil and natural gas for space heating or automotive transport. We hope other renewable energy sources will become more technologically refined and plentiful. As we develop renewable energy sources, electric power will inevitably become more important as a means of transporting power. Hydrogen as a means of storing energy may also become more important than it currently is. Hydrogen can be converted to electricity efficiently via a fuel cell.

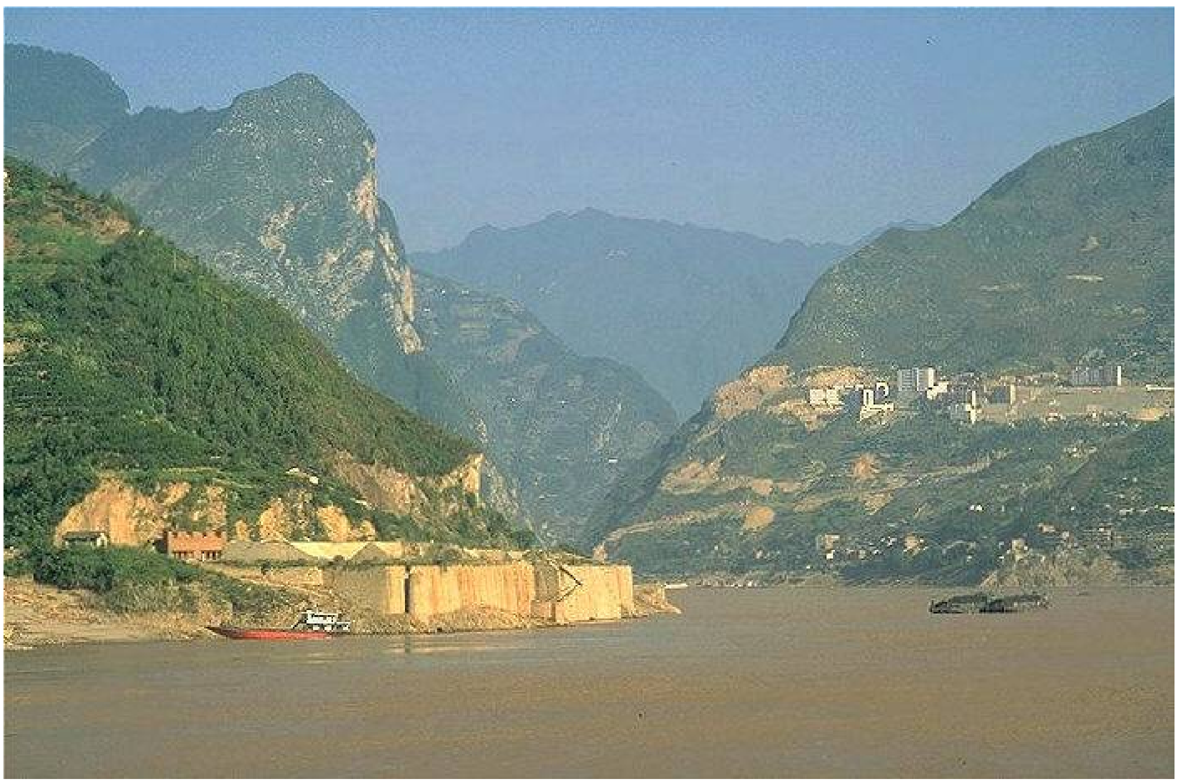


Figure 8. In cities along the Yangtze River entire cities were flooded and rebuilt up-hill to accommodate the water that was impounded behind the Three Gorges Dam. [9]

Just as the Niagara River in the area of the Falls has become a highly engineered environment, inevitably as time passes more of our environment will be subjected to development. An example of a modern and controversial hydroelectric project is the Three Gorges Dam on the Yangtze River in China. The potential of this area for hydroelectric power and flood control has been seriously studied and debated for nearly a century prior to the Chinese government's approval of the project in 1992. The reservoir for this dam eventually displaced over one million people from their homes and cities and raised the river level in three of the worlds most scenic river gorges. In return, flooding of the Yangtze River, which has already claimed over one million lives in the past century, will be controlled and about 22 GW of hydroelectric power will be available. The hydroelectric generation stations on the dam are scheduled for completion in 2009 and actually completed in about 2012. In matters small and large, engineers around the world participate in projects like this one.

6. AC CIRCUITS

The hydroelectric generators installed in the 1880's and 1890's at Niagara Falls were almost all alternating current (AC) generators. Prior to this date Thomas Edison had developed direct current (DC) lighting systems that had been installed in East Coast cities and elsewhere. There was a debate in the late 1800's over which type of system was best. The experience at Niagara Falls clinched the debate in favor of AC systems. At Niagara Falls transformers were used to step up the voltage from the generators prior to transmission. Then near the point of use of the electricity, other transformers were used to step the voltage down to safer levels again. High voltage transmission allows lighter and lower-cost over-head transmission lines, but the transformers only work with AC power. Although DC power has attractive features, such as a smooth flow of power using only one or two wires, without the possibility of using transformers (or something equivalent) the transmission of DC power was limited to a mile or two. Today, with modern electronics a device called a power inverter can be used to convert DC to AC. Then the voltage can be stepped up with a transformer. After transforming to a high voltage a rectifier can convert the AC back to high-voltage DC. After transmission a similar process can be used to step the high voltage DC back down to low voltage DC. Because of this possibility, DC power systems are enjoying a small resurgence of popularity for power transmission. Another important application for power inverters is with wind-powered generators. Some wind turbines generate DC power. Power inverters are then needed to connect to the AC power grid.

Definition: An electrical system is called a ***direct current (DC)*** system if the instantaneous current flow is always in the same direction. The current may be steady or pulsating.

Definition: An electrical system is a ***steady DC*** system if the amount of current flow is constant (or nearly constant) for long (usually a second or more) periods.

Definition: A DC electrical system that is not a steady DC system is a ***pulsating DC*** system.

Definition: An electrical system is called an ***alternating current (AC)*** system if it is not a DC system.

The characteristics of voltage do not necessarily follow those of current. For just one example, if the current is pulsating DC, the voltage does not have to be pulsating. The classification of AC or DC (alternating or direct *current*—note the word current) is made based on the direction of current flow alone. One can however speak of a steady voltage, pulsating voltage, or alternating voltage, with definitions similar to those for current.

A typical DC source is a battery. A typical AC source is an alternator. (An alternator is a coil of wire rotating in a stationary magnetic field or vice-versa, a stationary coil placed near a rotating magnetic field.) Pulsating DC is usually the result of converting AC to DC. The pulsations are sometimes undesirable. Then filter circuits are used to smooth the current flow to steady DC. Often in an AC circuit the waveforms of current and voltage are sinusoidal, but any waveform is theoretically possible. In today's more complicated circuits, non-sinusoidal waveforms are becoming more common.

Figure 9 illustrates the waveforms and resulting flows of power in some exemplary steady DC, pulsating DC and sinusoidal AC circuits. The left-hand column depicts the situation for a steady DC circuit, in this case a flashlight is used as a typical example. The voltage and current are constant, thus the instantaneous voltage, current, and power are also the average voltage, current and power. The power can be easily found from $P = VI$ using average (DC) values.

The center column of Figure 9 illustrates the power flow for a pulsating DC circuit. This example is the case of car battery charger that plugs into a wall outlet, steps the voltage down, and rectifies the current to pulsating DC. The voltages and currents shown are the battery voltage and the current flow into the battery's positive terminal (and out of the negative terminal). In this case it is true at every instant in time that $p(t) = v(t)i(t)$. The average power flow in this case is illustrated by the dashed line and the symbol P. Calculation of the amount of average power requires calculus but you can guess the average value of a waveform by making a judgment about areas. The average value is the level of power where the area of the power waveform above the average will exactly equal the area of the waveform below the average.

Definition: The ***average value*** of a waveform is that value for which the area of the waveform above the average value equals the area of the waveform below the average value.

The left-hand column of Figure 9 illustrates the situation for an AC circuit, in this case a motor. The voltage and current are given by the equations of sinusoids shown at the top of Figure 9. The average power is less than half the peak voltage times the peak current.

In the example of the motor (Figure 9) note that the average AC voltage is zero. The average AC current is also zero. The power at any instant in time is the product of the instantaneous voltage and instantaneous current. Sometimes the power is negative. At those times the motor is actually generating electrical power. Just as a flywheel can deliver power to a mechanical load for a short time, so can a motor act like a generator for a short time, getting the needed energy from the inertia of its rotating parts and from changing magnetic fields. (Motors are not the only electrical devices that can momentarily act like generators.) Most of the time the power is positive, meaning the motor is absorbing electrical power. The average value of the electrical power, P, illustrated by the dashed line, is positive in this case, not zero as the equation $P = VI$ (using average values) would predict. Also observe from the graph for the power, $p(t)$, that electrical power pulsates at twice the frequency of the voltage and current.

The electrical power consumed by the motor is turned into something else that is useful, such as the rotation of the motor's shaft to power a mechanical load. The inertia of the rotating parts of the motor (and mechanical load) is helpful to smooth out the flow of power and push the rotation of the shaft through the times when the electrical power flow is negative. The pulsation of the electrical power flow will cause the motor to vibrate and hum. (Poor mechanical balance and mechanically loose parts that can wiggle in the changing magnetic fields can also cause vibration and hum.) The mechanical power output of the motor will be a small amount less than the average electrical power, depending on the efficiency of the motor. Most AC motors are at least 80% efficient, and 95% efficiency is not too hard to achieve. Electrical power not converted into mechanical power is converted to heat.

AC Voltage $v(t) = 165\sin(120\pi t)$ V

AC Current $i(t) = 1.41\sin(120\pi t - \pi/4 \text{ radians})$ A

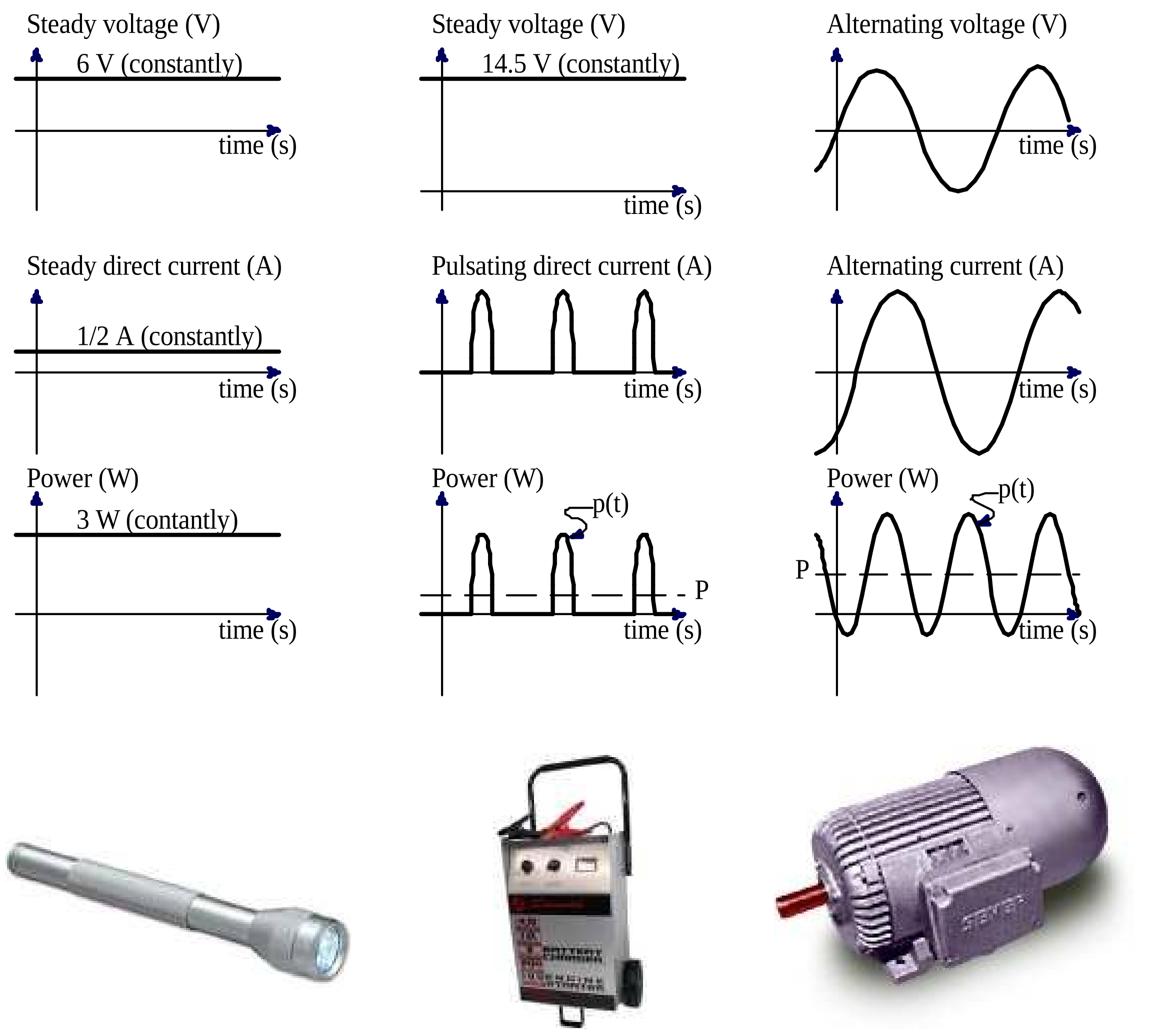


Figure 9. DC and AC voltages, currents, and resulting power flow in three situations.

The actual formula for the AC power flow when the voltage and current waveforms are sinusoidal can be found by multiplication of the *instantaneous* voltage times the *instantaneous* current. Some further manipulation with trigonometric identities can be used to find the average power, in this case without the need for calculus.

$$p(t) = v(t)i(t) = [165\sin(120\pi t)][1.41\sin(120\pi t - \pi/4 \text{ radians})]$$

$$p(t) = 233[\sin(120\pi t)][\sin(120\pi t - \pi/4 \text{ radians})]$$

Trig ID: $\sin(\alpha)\sin(\beta) = 1/2\{\cos[\alpha - \beta] - \cos[\alpha + \beta]\}$,
where in this case let $\alpha = 120\pi t$ and $\beta = (120\pi t - \pi/4 \text{ radians})$

$$p(t) = (233/2)\{\cos[120\pi t - (120\pi t - \pi/4 \text{ radians})] - \cos(120\pi t + (120\pi t - \pi/4 \text{ radians})]\}$$

$$p(t) = 116[\cos(\pi/4 \text{ radians}) - \cos(240\pi t - \pi/4 \text{ radians})]$$

$$p(t) = 82.3 - 116\cos(240\pi t - \pi/4 \text{ radians})] \qquad \textbf{(Equation 1)}$$

The first term of Equation 1 (above) shows that the average power is 82.3 watts. The second term averages to zero, but it shows that the power is pulsating and it gives the peak magnitude of the pulsations (±116 W).

Recall that the instantaneous amount of electric power is found from the formula $p = vi$, or more completely, $p(t) = v(t)i(t)$. These two formulae are always true at any instant in time. In a DC circuit we can also write $P = \bar{V}\bar{I}$ since the average and instantaneous amounts are equal ($\bar{V} = v$, $\bar{I} = i$ and $P = p$), but $P = \bar{V}\bar{I}$ (using average values) only applies to steady DC circuits. For an AC circuit or a pulsating DC circuit, $P = \bar{V}\bar{I}$ is almost never true. In the above example for the motor $\bar{V} = 0$ and $\bar{I} = 0$ but $P = 82.3$ W.

Since the average amounts of voltages and currents in AC circuits are usually zero, we need some other way to describe voltage and current magnitudes in AC circuits. The analysis of this problem relies on calculus and goes beyond the scope of this course, but the results of the analysis can be applied easily if the waveforms are sinusoidal. If the peak magnitude of a sinusoidal voltage or current is divided by the square root of two, the result is the value of a steady DC voltage or current that has the same ability to deliver power. This is called the *root-mean-square* (RMS) amount or sometimes it is called the *effective value*. That this magnitude of a sinusoidal voltage or current has the ability to deliver the same amount of power as that amount of steady DC voltage or current has does not mean that it necessarily will do that in practice. Much depends on the characteristics of the AC circuit.

Definition: The ***Root Mean Square*** value of any current or voltage waveform (pulsating or alternating) is the value of a steady DC amount that has the capability to deliver the same amount of power. Abbreviated ***RMS***. Also known as *effective value*.

Theorem: The RMS value of a sinusoidal current or voltage waveform is the peak value of the waveform divided by the square root of two.

In most situations involving AC systems, RMS amounts are used for voltage and current without comment or explanation. When we say that, "the standard power line voltage in North America is 117 V" that voltage is in units of volts RMS. Likewise, when you switch a meter to AC mode, the measurements become RMS values. (Power is given in instantaneous or average amounts, there are only a few rare cases where RMS amounts of power have any meaning—and these have nothing to do with the definition given above since that definition relates only to voltage and current, not the product of voltage and current which is power.)

Some meters for AC current and voltage will give bogus readings if the waveshape is not sinusoidal. These meters assume the waveform is sinusoidal and cannot respond correctly to changes in the waveshape. Other meters (more expensive ones), when switched into AC mode will show the correct RMS value for any waveshape, including pulsating direct currents and voltages. These meters are called "true RMS" meters.

In the right-hand column of Figure 9 the peak value of the voltage is 165 V. Dividing by the square root of two gives 117 V RMS. The peak current is 1.41 A and dividing by the square root of two gives 1.00 A RMS. If RMS voltage and RMS current are multiplied together you get the maximum possible average power, 117 W in this case. The actual average power usually is less than that, 82.3 W in this case. This is because the peaks of the voltage do not align with the peaks of the current, as illustrated in Figure 9. When the peaks do not align perfectly, which is common, then there are instants in time when the power is flowing backwards, which reduces the average power flow below what is possible. This gives rise to the definitions of *apparent power*, *average* (or *real*) *power*, and *power factor*.

Definition: In an AC circuit, the ***apparent power*** is the maximum possible power that can flow based on current and voltage. It is symbolized by |S|. It can always be found as $|S| = (V_{rms})(I_{rms})$

If the waveforms are sinusoidal and the voltage and current are given in RMS quantities, then generalizing the computation of the first term from Equation 1 we conclude that in an AC circuit the average power can be found as follows:

$$P = (1/2)V_{peak}I_{peak}\cos(\alpha - \beta)$$

$$P = \left(\frac{V_{peak}}{\sqrt{2}}\right)\left(\frac{I_{peak}}{\sqrt{2}}\right)\cos(\theta)$$

$$P = V_{rms}I_{rms}\cos(\theta) \qquad \text{(valid for sinusoidal AC circuits)}$$

where θ (theta) is the phase lag (in radians or degrees) of the current waveform measured with respect to the voltage waveform. In Figure 9 that phase lag is $\pi/4$ radians. (If the current waveform is leading rather than lagging the voltage waveform then we say the phase lag is a negative amount.)

Definition: In an AC circuit, the ***power factor*** is the ratio of the average power flow to the apparent power flow. It is symbolized by *pf*. It can always be found as $pf = \cos(\theta)$.

Thus for an AC circuit with sinusoidal wave shapes, the average power flow is found as:

$$P = V_{rms} I_{rms} (pf) \qquad \textbf{(Equation 2)}$$

As far as getting useful work done is concerned, it is the average power, *P*, that matters. Usually the net power factor of a whole household or small industry is about 0.8 or higher. The electric companies just factor this slight issue into their price and bill according to the real energy used. Utility companies are concerned however that the power factor does not get too low because an ordinary electric meter responds to average power, not apparent power. Loads with a low power factor, such as a lightly loaded electric motor, require higher than necessary current (to offset the low power factor) and put unnecessary stress on the utility company. Large consumers of electricity might be billed a surcharge if their power factor is too low.

Three-phase AC electric power is a type of AC power that allows three conductors to work together to deliver power in a constant flow with no pulsations, just like DC circuits do. An inter-city three-phase power transmission line usually has main conductors made of a stranded steel wire-rope core for strength and a surrounding layer of aluminum strands for conduction. The whole bundle of strands for one conductor can be as large as about two inches in diameter. These large conductors can carry thousands of amperes of current. High voltage levels are also typically used for electric power transmission. Inter-city three-phase power lines typically operate at up to three-quarters of a million volts. Such three-phase transmission lines can carry very large amounts of power, of the order of a gigawatt.

Example Set #3

The label on a certain bread toaster designed for use in a typical North American home shows that it consumes 1000 W. Assuming that the phase lag of the current with respect to the voltage is zero radians ($\theta = 0$), figure out how much current the toaster draws.

We assume also that the AC power line voltage is 117 V RMS and that the waveforms of the voltage and current are both sinusoidal. Then $P = V_{rms} I_{rms} \cos(\theta)$. Solving for I_{RMS} gives

$$I_{RMS} = P/(V_{RMS}\cos(\theta))$$

$$I_{RMS} = (1000\text{ W})/[(117\text{ V})(\cos(0)]$$

$$I_{RMS} = (1000\text{ W})/[(117\text{ V})(1)] = 8.55\text{ A}$$

We observe that in order to carry this much current, the power wire for the toaster is made of a heavy gauge of stranded copper.

Suppose that the above toaster is to be redesigned for use in a camper-trailer that uses a 12.6 V car battery for electric power. How much current will the toaster have to draw in order to operate at the same power level?

For a steady DC circuit $P = VI$ and we are given that $P = 1000$ W and $V = 12.6$ V. A little algebra gives

$$I = P/V = (1000\text{ W})/(12.6\text{ V}) = 79.4\text{ A}$$

We observe that this idea for a toaster for a camper-trailer is not going to be very practical due to the high current needed. The power wiring connecting to the toaster to the battery would have to be as heavy as a pair of jumper cables typically used to start cars with dead batteries. Such heavy wiring would cost more than the toaster.

Suppose the toaster is to be redesigned for the European market where the normal power line voltage is 240 V RMS. How much current will the toaster draw then?

$$I_{RMS} = P/(V_{RMS}\cos(\theta))$$

$$I_{RMS} = (1000\text{ W})/[(240\text{ V})(\cos(0)]$$

$$I_{RMS} = (1000\text{ W})/[(240\text{ V})(1)] = 4.17\text{ A}$$

We observe that the line cord for the European toaster does not need to be as heavy as the one on a North American toaster, but the insulation must be better. We further observe that since insulation is usually cheaper than copper, higher voltage circuits are usually more economical when large amounts of power are needed.

Example Set #4

How long can a certain DC power transmission line be, how much power can it deliver, how much power will it waste, and how much will it cost? A single conductor of this transmission line is 1 cm in diameter and it contains mostly aluminum. The whole transmission line costs $30 per meter of length. (This includes the cost of land, towers, construction labor, etc.) Suppose that there will be a voltage drop of 0.1 volt per meter when the current is 300 A, and that 300 A is the maximum allowable current. Also suppose that a DC generator can produce at most 10 kV and the load at the end of the transmission line needs at least 5 kV.

Let the current be the maximum, 300 A. Also, to maximize the length of the transmission line, maximize the total end-to-end voltage drop in the line subject to maintaining the load voltage at 5 kV. Then the power delivered to the load is

$$P = VI = (5\text{ kV})(300\text{ A}) = 1.5\text{ MW}$$

Since the transmission line requires two wires, one to bring the current to the load and one for the return circuit, the voltage drop allowed on each wire is 2.5 kV. (2.5 kV + 2.5 kV + 5 kV = 10 kV, the voltage drops add up to the voltage at the generator.) Now extend the length of the transmission line until the voltage drop in a conductor is 2.5 kV

$$X = (2.5\text{ kV})/(0.1\text{ V/m}) = 25\text{ km}$$

The length, 25 km, is about 15.5 miles. The cost is

$$(25\text{ km})(\$30/\text{m}) = 750\,000\text{ dollars.}$$

The power wasted by the transmission line will be

$$\text{P} = (2\text{ conductors})(2.5\text{ kV})(300\text{ A}) = 1.5\text{ MW}$$

The power that the generator puts out will be

$$P = VI = (10\text{ kV})(300\text{ A}) = 3\text{ MW}$$

Note that this very expensive transmission line wastes half of the power generated.

Now suppose that the same conductor is used with AC current and transformers are installed on each end to step the voltage on the transmission line up to 500 kV. Also suppose the voltage drop on the transmission lines will be only 1 % of the source-end voltage, or 2.5 kV per conductor. (All voltages and currents are RMS amounts now since we are working with an AC system and sinusoidal wave shapes.)

The voltage at the load end will be

$$500\text{ kV} – 2.5\text{ kV} – 2.5\text{ kV} = 495\text{ kV}$$

Now the maximum power that *could* be delivered to the load (assuming the voltage and current are exactly in phase with each other) is:

$$P = V_{rms}I_{rms}\cos(0) = (495\text{ kV})(300\text{ A}) = 49500\text{ kW} = 148.5\text{ MW}$$

The length of the line and power lost will be the same as before since the voltage drop on each conductor remains 2.5 kV and the current remains 300 A. The cost will be a few hundred thousand more to cover the cost of a transformer at each end. Using higher voltage on the transmission line made it possible to deliver about 100 times more power and improved the efficiency to about 99%. (The analysis of long transmission lines is much more complicated than this simple example of a short transmission line.)

Further study of AC circuits is beyond the scope of this course. From this point forward we will concentrate on DC circuits and signal processing.

7. RESISTORS

In example set #4 the voltage drop was stated to be 0.1 volts per meter when the current is 300 A. This loss of voltage (loss of potential to do work) is a phenomenon called resistance.

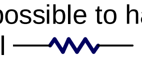
Definition: A material's opposition to electric current is called electrical ***resistance***. It is quantified as the ratio of voltage to current. Symbol *R*. Mathematically, $R = v/i$

Definition: The ***ohm*** is the SI unit for resistance. A material that has a one volt drop between two points when there is one ampere of current flowing from the positive point through the material to the negative point has a resistance of one ohm between the two points. Abbreviated as "Ω" (capital omega).

The equation $R = v / i$ is known as *Ohm's law*. (The equations $v = Ri$ or $i = v / R$ are also called Ohm's law since they are algebraically equivalent.) Interestingly the equation is not the claim that Georg Ohm made. Ohm claimed that the resistance of a piece of metal held at a constant temperature is a constant.[10] He had to use the equation in order to make that claim. Although the equation is memorable, it is only incidental to his claim that metals at a constant temperature have constant resistance. Any material for which *R* is practically a constant is called an *ohmic* material. Many important materials, such as doped semiconductors (silicon, germanium, etc.) are not ohmic. Modern electronic devices absolutely depend on non-ohmic materials precisely because resistance in these materials is not a constant but can vary with voltage and current. It would be more honorable to the exact words of Georg Ohm if we said that $R = v / i$ is the "definition of resistance," that he invented this definition, and that "Ohm's law" states that metals have constant resistance regardless of the amount of voltage or current. Unfortunately that is not the way the matter has developed. "Ohm's Law" is understood to be $R = v / i$, the mathematical statement of the definition of resistance.

The cause of electrical resistance has to do with the interactions of the charge carriers (often electrons) and the atomic structure of the material. One could envision that the more atoms an electron bumps into as it travels through the molecular (or crystal) structure of the material, the higher the resistance will be. Further development of a physical understanding of the causes of resistance is a typical subject in college physics courses.

In some cases, such as in the case of an electric power line, resistance is something to be minimized. In other cases, such as when we deliberately want to make heat or moderate or control the current, resistance is a good thing in a controlled amount. When we build a device deliberately to control or moderate the flow of current in an ohmic (constant resistance) way, we call the device a resistor.

Definition: A ***resistor*** is a two-terminal device designed as far as possible to have a constant resistance between the two terminals. Symbol 

If a resistor has an adjustable amount of resistance, as for example a resistor with a sliding connection, it may be symbolized with an arrow drawn through it. Some adjustable resistors have three terminals. The middle terminal represents a slider that can be moved or rotated. This three-terminal style of adjustable resistor is called a *potentiometer*. Its symbol shows an arrow pointed at the resistor to represent the adjustable slider.

Figure 10. Symbols for adjustable resistors.

Example Set #5

What is the resistance of the bread toaster designed for North American home use? (see example set #3)

$$R = V/I = (117\ \text{V})/(8.55\ \text{A}) = 13.7\ \Omega$$

What is the resistance of the bread toaster designed for the camper-trailer?

$$R = V/I = (12.6\text{ V})/(79.4\text{ A}) = 0.159\ \Omega$$

What is the resistance of the DC transmission line in example set #4?

For each wire

$$R = V/I = (2.5\text{ kV})/(300\text{ A}) = 8.33\ \Omega$$

Since there are two wires, the total resistance is 16.7 Ω.

What is the resistance of the flashlight illustrated in the first column of Figure 9?

$$R = V/I = (6\text{ V})/(0.5\text{ A}) = 12\ \Omega$$

8. CIRCUIT ELEMENTS: MATHEMATICAL MODELS, AND THE PASSIVE SIGN CONVENTION

In order to do engineering design work we need a way to predict the behavior of real-world electrical devices, like the resistors discussed above. We create mathematical models of the devices in order to predict their actual behavior. For electrical devices we call these mathematical models *circuit elements*. Circuit elements are ideal, whereas the corresponding real-world devices are never as perfect.

Definition: A ***circuit element*** is a mathematical model used to describe a class of actual electrical devices.

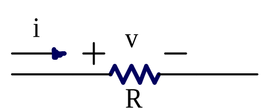
Definition: A ***simple circuit element*** has these characteristics

1.) It has exactly two terminals
2.) There exists a known mathematical relationship between the voltage across and the current through the circuit element.
3.) The circuit element is irreducible to (cannot be equivalently represented by) another simple circuit element or a combination of *other* simple circuit elements.

In order to give the algebraic signs in mathematical equations valid meaning, we need to somehow assign directions to current flows and voltage polarities in circuit diagrams. This is done by means of an agreed upon convention. By convention the defining mathematical relationship for a circuit element is valid when the current *arrow* is directed into the positive *labeled* terminal.

Definition: The ***passive sign convention*** is that the current arrow shall be directed into the positive labeled terminal of a circuit element.

The passive sign convention for a resistor can be illustrated as shown in Figure 11. Note how the current arrow is directed into the positive labeled terminal.



$v = Ri$

Figure 11. A circuit element model includes three things:
1.) The symbol for the electrical device, in this example a resistor.
2.) Labels that follow the passive sign convention, and
3.) An equation relating voltage to current.

The passive sign convention is only about labels written on paper. It does not say that the current will necessarily flow into the positive terminal. This is because the amount of voltage and/or current can be negative. Consider the case if v = –20 V. and R = 5 Ω. Then according to the equation that models a resistor ("Ohm's Law") $i = v/R$ =(–20 V)/(5 Ω) = –4 A. In other words, in this case the actual net positive current flow is opposite to the direction of the arrow on the paper and the actual voltage polarity is opposite to the labels on the paper.

It is important to note that a mathematical *variable* for current or voltage *must have a label* (an arrow or a + – pair) on a schematic to give it a non-ambiguous meaning.

The passive sign convention is vital to using mathematics to calculate actual current directions and voltage polarities. When the schematic is initially labeled, the actual directions of current flow and polarity are sometimes unknown. This is not a problem. Just label the schematic diagram with arrows and polarity marks that are in agreement with the passive sign convention. Then solve the problem using those labels and the equations from the circuit element models. The sign of the answer tells us if the labels actually point in the right directions. Usually if the quantities found are negative we do not bother to change the labels on the schematic and recalculate the answers. We just specify a negative number for the answer and use the labels on the schematic as originally written. Note that the situations illustrated below in Figure 12 and Figure 13 all represent exactly the same situation.

Figure 12. An illustration of how the passive sign convention is used. In both situations the flow of net positive current is from right to left. In both situations the right side of the resistor has the greater voltage in comparison to the left side.

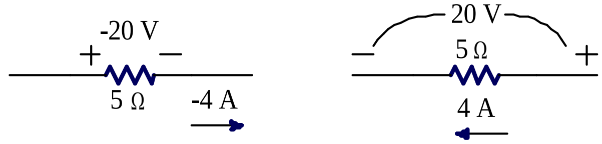


Figure 13. Labels may be slid around on the page but not rotated. Both of these cases satisfy the passive sign convention in that the current arrow is effectively pointed into the positive labeled terminal. They again illustrate the same situation depicted in Figure 12

The passive sign convention has a valuable additional benefit. It can be used to show the direction of instantaneous electrical power flow or the direction of average power flow in a steady DC circuit. When the passive sign convention is used, a positive amount of power means the circuit element is absorbing electrical power (and converting it to some other useful form of power). When the passive sign convention is used and the power calculated is negative, then the circuit element is supplying electrical power to the circuit.

Example Set #6

How much electrical power does the starter motor absorb in the situation of example set #1? How much electrical power does the battery supply? Recall that the battery voltage is 12 V and the starter draws 100 A.

First we draw a schematic of the situation and add labels for each circuit element to satisfy the passive sign convention.

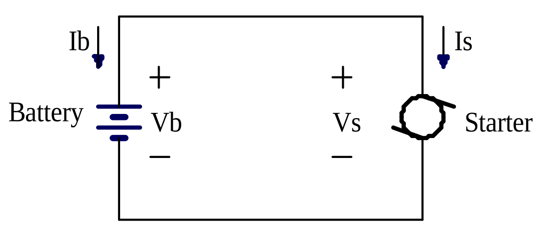


Figure 14. Labels (*I_b, I_s, V_b, V_s*) with polarity marks (arrows for current, a + – pair for voltage) are important parts of any schematic. This is a schematic of a car's starter circuit.

Comparing the schematic to the given information we conclude that V_b = 12 V and I_s = 100 A. Since the battery and starter are connected to each other, $V_s = V_b$ = 12 V. Since I_b is directed opposite to I_s we conclude that $I_b = –I_s$ = –100 A. Now calculating the electrical power absorbed by the starter we get:

$$P_s = V_sI_s = (12\text{ V})(100\text{ A}) = 1200\text{ W} = 1.2\text{ kW}$$

A similar calculation for the battery gives:

$$P_b = V_b I_b = (12\text{ V})(–100\text{ A}) = –1200\text{ W} = –1.2\text{ kW}$$

This shows that the battery is supplying +1.2 kW of electrical power (or absorbing –1.2 kW) and the starter is absorbing 1.2 kW of electrical power.

Many synonymous words can be used in describing the flow of power. "Delivering," "giving," "producing," etc. all mean the same thing as "supplying" in this context. Similarly, "using," "drawing," "consuming," etc. all mean the same thing as "absorbing."

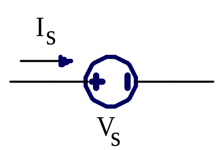
Some students find the passive sign convention counterintuitive when it is applied to a power source such as the battery illustrated in Figure 14. Their argument is that current actually tends to flow out of the positive terminal of a battery. Why then do we place the current arrow for the battery, I_b,, "backwards?" The answer is that an assumption that current flows *out* of a positive terminal of a battery (or out of any positively labeled terminal) is not universally reliable. When a battery is charging, then the net positive current flow is *into* the positive terminal. The idea of labeling a schematic according to the directions we think current will flow, rather than according to the passive sign convention, relies on understanding the circuit prior to doing the analysis! In complicated situations how can we really know which direction current will flow until we do some analysis? The passive sign convention provides a reliable means to mathematically calculate the actual directions.

The model of a resistor is not the only simple circuit element. Other simple circuit elements are independent voltage sources, independent current sources, capacitors, and inductors. Capacitors and inductors involve concepts and differential calculus that goes beyond the scope of this course, but voltage sources and current sources are easy to understand.

Definition: An ***independent voltage source*** is a two-terminal circuit element (a model) that maintains a constant voltage between its terminals regardless of the current flow through it.

There are such things as dependent voltage sources, but these are a topic for a more advanced course. Usually the word "independent" is omitted—just implied. That will be the practice in this text from here on. If a source is "dependent," then the word "dependent" must always be stated to avoid confusion. Assume a source is "independent" unless otherwise stated.

A voltage source can be used to model a battery, for just one possible example. You might argue that there is no relationship between the voltage and the current since the current can be anything. Mathematicians consider it the other way around however. Given the current, you can use the model to predict what the voltage will be. Although you do not really need to know the current, you can predict the voltage—that is a relationship of sorts! This is called a degenerate relationship. (Another example of a degenerate relationship is to note that a circle of zero radius degenerates to a point. A point is a degenerate circle.)



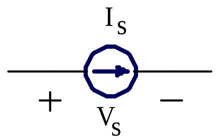
V_S = (a constant)

Figure 15. The symbol and equation for a voltage source [an independent voltage soucre].

Another important simple circuit element is an *independent current source*. The relationship between voltage and current is that the current is constant regardless of voltage. (Another degenerate relationship, but given the voltage, you can figure out the current!)

Definition: An ***independent current source*** is a two-terminal circuit element (a model) that maintains a constant current through it regardless of the voltage across its terminals.

There are few parts in electrical circuits that can be modeled by a current source alone. Perhaps one possible example is the coil in a car's ignition system. It forces a current through the gap of a spark plug although there is no wire or other obvious conductive path for the current. The coil creates whatever voltage is necessary (a high voltage) to ionize the air-fuel mixture in the spark gap. The electric field created by the high voltage across the gap is strong enough to pull the electrons out of their normal orbits and away from the protons in the nucleus of the atom. As a result the electrons (negatively charged) and the remainder of each atom stripped of some electrons, now called an ion (and positively charged) are free to travel in opposite directions and the air-fuel mixture becomes electrically conductive. The energy involved in this process also produces a flash of light and heat—a spark. In the sense that the coil forces a current to flow no matter what, the coil could be modeled as a current source. However, a coil is not a perfect current source. If the gap in the sparkplug is too long then the resulting electric field will not be concentrated enough to produce ionization and the current will not flow.



I_S = (a constant)

Figure 16. The symbol and equation for a current source [an independent current source].

In summary, three types of simple circuit elements have been introduced: a resistor, an independent voltage source (usually known simply as a voltage source) and an independent current source (or usually called a current source). In order to make sense of the directions of current flow and voltage polarities used with these models the passive sign convention must be used. In later courses in electrical engineering you can learn about two more simple circuit elements, capacitors and inductors. There are also non-simple circuit elements such as potentiometers, transformers, dependent sources, and transistors, which are introduced in other courses.

9. HISTORY: POWER SYSTEMS, AND SIGNAL PROCESSING

So far this text has focused on the subject of electrical power systems. Electrical Engineering has traditionally included two other sub-disciplines. In addition to *Power Systems*, there are *Communication Systems* (or analog electronics) and *Computer Engineering* (or digital electronics). However in recent years, the last two sub-disciplines have merged into one, called *Signal Processing*. In another decade or two as power systems become more complex, the divisions between sub-disciplines in electrical engineering might become entirely unimportant.

The development of power systems engineering as a recognized discipline started in the 1880's with the development of electrical lighting, generators, and motors. Well known pioneers in this work include Thomas Edison, Nicola Tesla and George Westinghouse. The hydroelectric power stations at Niagara Falls, being the first large-scale developments in power systems engineering, brought public attention to the discipline. This motivated colleges and universities to introduce electrical engineering majors into their curricula and the field became a recognized area of study. Those working in this field desired a society in which to discuss these matters. In 1884 The American Institute for Electrical Engineers (AIEE) was founded in New York City. The interest of the AIEE was light and power, but this expanded to include wired telegraph and telephone systems.

Communication systems (meaning radio and analog electronics) developed starting at about the same time, around 1890, with the work of Guglielmo Marconi. Starting in about the 1920's electrical systems were increasingly used to control industrial processes and military equipment. The theory of electrical remote-control systems was derived in large part from the theory of communication systems. These developments prompted further specialization in the engineering courses in universities and recognition of the sub-discipline of communications and/or control systems. The people working in communications and control systems also wanted a society in which to discuss matters. In 1912 The Institute of Radio Engineers (IRE) was founded via the merger of two smaller regional societies (Society of Wireless Telegraph Engineers and The Wireless Institute) and started gaining an international membership. The IRE became the leading professional society promoting and documenting the technical development of radio communication and remote-control systems.

Computer engineering developed starting around 1940, motivated in part by the work done during World War Two by various governments to crack each other's diplomatic and military codes. During the 1980's and 1990's however, communications, controls, and computers have merged together as a sub-discipline, now called *Signal Processing* (or sometimes Information Processing, but that phrase is falling out of style). This merger happened because formerly analog systems, like the telephone, radio, phonograph, and television were replaced by what at least in layman's terms appear to be digital equivalents. These are cellular telephones, satellite radio, CD's and DVD's, and high-definition television. The design of these new devices combine digital signals with many techniques first developed for analog electronics such as sophisticated modulation and coding theories. Since most electronic devices, for one example telephones, now have little computers in them and since computers can act like other devices, for example computers can act like telephones, there is no longer a clear bright line that distinguishes a computing device from a communication device. Similarly, some things that appear to be digital, like a cellular telephone, contain sophisticated analog circuits, such as a transmitter, and so there is also no distinct division between an analog device and a digital device anymore.

In modern electrical engineering designs one finds the traditional divisions between sub-disciplines becoming ever less important and general engineering knowledge and system-level thinking becoming more important. For example electric power companies are currently experimenting with methods to deliver a broadband digital connection (Internet) to every power outlet. Also, the control of modern power delivery systems now relies on sophisticated computer systems and satellite communications. Even the division between power systems and signal processing is not as sharp as it once was.

The intertwining of these sub-disciplines was already recognized by leading engineers in the early 1960's. They proposed a merger, and in 1963 the AIEE and the IRE merged to form the Institute for Electrical and Electronics Engineers (IEEE) [11]. The IEEE now has a worldwide membership of about 365000 members including about 65000 students. Although the IEEE was founded and initially grew mostly in the United States, it now has a decidedly international flavor and most of its growth is occurring outside the North American region. A similar society, The Institution of Engineering and Technology (IET) can trace its origins back to 1871 in England. It has about 150000 members [12].

10. "CQD"

Wireless telegraphy was a relatively new phenomenon at the time the *Titanic* sank. The *Titanic* was equipped with a state-of-the-art wireless system. This was installed as a business venture to make money by sending personal messages on behalf of the passengers. Telegraph messages were charged by the word at that time. These were called "commercial" messages. The wireless systems were also used to coordinate shipping operations, such as to facilitate planning for arrival at a harbor. Considering the intended primary purpose of the wireless, most wireless systems were operated only during daytime hours when passengers wanted service and when ports and shipping companies were open for operations. (The range of communication was better at night, but few people wanted to communicate via radio at night.)

On the night of April 14, 1912, a Sunday evening, at about 11:40 PM Harold Cottam, the radio operator on the ship *Carpathia*, was at the end of his shift. His captain had stopped *Carpathia* dead in the water for the night since an ice field had been spotted ahead. The plan was to resume travel in the morning when lookouts could see better.

> [Cottam] was undressing before retiring, but still wore his headphones, hoping to catch a reply to a message he had earlier dispatched. Still not wanting to give in yet, he twiddled with the tuner, coming across *Titanic's* unusually quiet commercial frequency. He tapped out a message to the liner, "I say old man, do you know that there is a batch of messages coming through for you from MCC?" ["MCC"—Marconi Cape Cod Shore Station] Almost before Cottam finished the transmission, Phillips [Titanic's radio operator], came back with "CQD." ["CQD"—Come Quick. Danger!" It was then the distress signal of the most urgent kind.] Cottam, who was taken aback by this monumental transmission, double checked with Phillips, "Shall I tell my captain? Do you require assistance?" To which Phillips replied, "Yes. Come quick!" [13]

The *Carpathia* was about 70 miles from the *Titanic* at the time of this exchange of messages. At about the same time, the *Californian* was only about 10 miles from the Titanic. Officers on the *Californian* saw distress flares from the *Titanic* but mistook them for a fireworks show. The *Californian* tried to contact the *Titanic* by signal light, but those on the *Titanic* never saw the signal lights from the *Californian*, or even knew it was in the area. The *Californian's* radio operator had retired for the night and the captain of the *Californian* did not think of trying the wireless since it was considered mainly for commercial messages.

Figure 17. Jack Phillips, *Titanic's* radio operator (left), and his deputy, Harold Bride. [14]

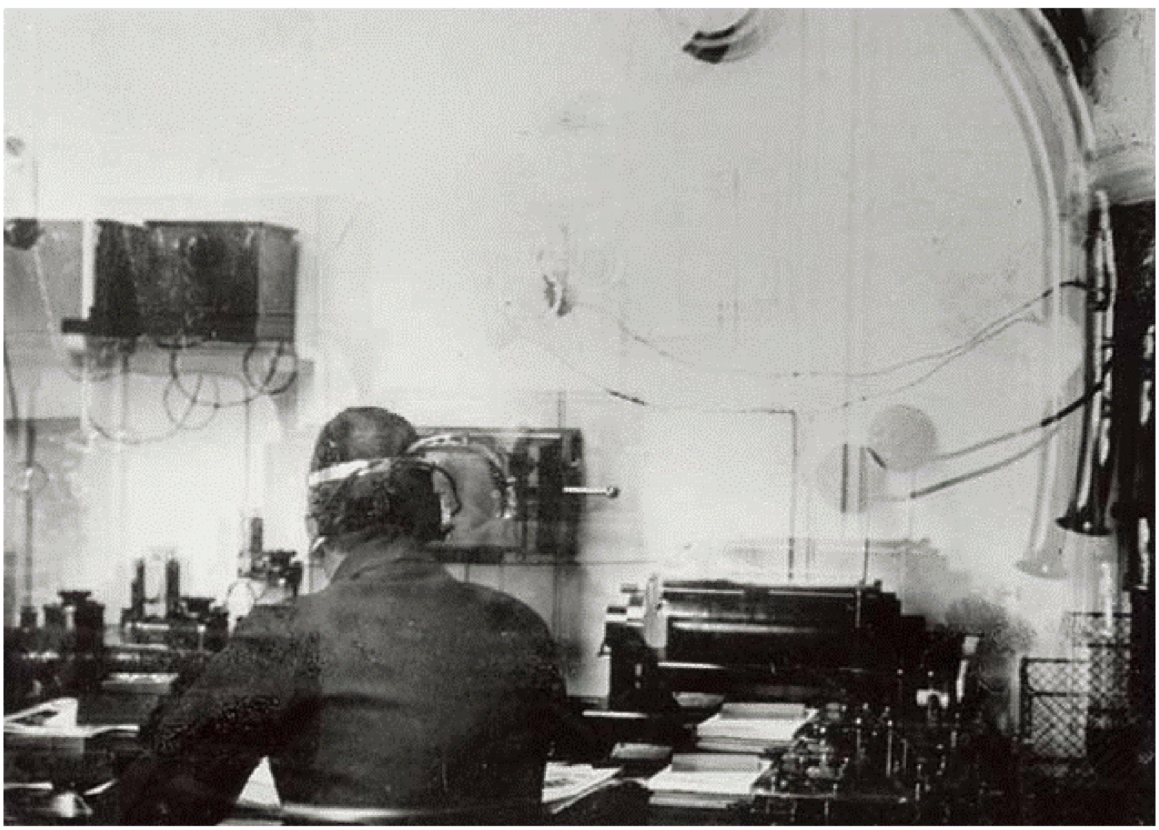


Figure 2, The only known photograph of the *Titanic's* Wireless Room. Taken by passenger Fr. Browne. The operator is believed to be Harold Bride. [14]

The role of the wireless in rescuing 705 of *Titanic's* passengers catapulted the technology to a new level of public awareness and consequently, engineering development. The present day discipline of signal processing continues to contribute to public safety now with automatic distress buoys linked to the global positioning system, Doppler weather radar, and in countless other ways.

11. BANDWIDTH AND THE SHANNON HARTLEY- THEOREM

Commercial AM radio broadcasting—the type of AM radio we are familiar with—started in November, 1920, when radio station KDKA went on the air in Pittsburgh. The station's first program was news of election results. Soon after, other stations went on the air. Generally there could be only one station per city. At that time all commercial broadcast stations were on the same frequency, 833 kHz. This was done because early radio receivers did not have the types of tuning dials we are accustomed too. Adjusting the tuning was too difficult to be practical. Later, there was competition within cities as more businesses desired to put radio stations on the air. This was one of the first episodes of a need for more bandwidth! At first stations in the same city cooperated with each other by a scheme called time-sharing. Only one station at a time would be on the air and they would take turns. Since all broadcasts were live at this time, the time off-the-air gave stations a chance to plan and arrange for their next program. So for example at 5 PM you could hear the news on a station operated by, and located at a newspaper. Then, at 6 PM that station would go off the air and another, operated by and located at a nightclub, for example, would go on the air with music, etc. Most stations operated as a promotion for an underlying business. [15]

Time-sharing did not last long. Beginning in 1923 the AM band was enlarged to 550 – 1350 kHz. A year later the top end was raised to 1500 kHz. In 1941 the top end was raised again to 1600 kHz [16], and in 1997 the top end was again raised to 1710 kHz. [17] The era from 1920 to about 1950 is known by many in the broadcasting business as the "golden age" of radio since at that time AM radio was the only broadcast medium that could effectively reach the general public. Today, with the advent of digital broadcasting there continues to be a tremendous demand for access to the radio spectrum for various radio-based services such as cellular telephones, high-definition television, global positioning systems, and so forth. In one era the public wanted lots of radio and television stations to be available. The public still wants lots of those, but now they also want faster Internet service and better quality cellular telephone connections, etc. too. How much information can be sent by radio?

Theorem: ***The Shannon-Hartley Capacity Theorem***

The capacity of a communication channel (wired or wireless) is

$$C = B \times \log_2(1 + S/N), \text{ where}$$

C is the channel capacity in bits per second.
B is the channel's bandwidth in Hz.
S is the signal power in W
N is the power of Gaussian noise in the channel, in W

Furthermore, if the needed information rate in bits per second, R, is less than C, then there exists a coding technique that will allow the information to be transmitted through the channel with an arbitrarily small number of errors. On the other hand, if C is less than R, then the error rate will increase without bound.

For a long time the question of how much information could be broadcast in a given channel of bandwidth was an open question. In 1949 Claude Shannon published his master's degree thesis which included an answer to the question. His most famous theorem is named the Shannon-Hartley Capacity Theorem in honor of Ralph Hartley, an engineer at Bell Labs who made many contributions to information theory. An especially surprising part of the theorem is the claim that if $R < C$ then there exists a coding method which gives practically error free communication. This was an amazing claim for its day since error correction codes were only in their infancy in 1949. Unfortunately, the theorem does not say how to code the information in order to achieve the performance the theorem proves is theoretically possible.

According to the Shannon-Hartley Capacity Theorem, communication channels of the 1950's era were effectively transmitting only a tiny fraction of the information rate the theorem predicted was possible. These communication systems remain in widespread use for AM and FM broadcasting, analog television broadcasting, and "plain old telephone" (wired) service. It took about 40 years to figure out how to code information to realize the rates the theorem predicted. In the 1990's engineers perfected a technique called "combined coding and modulation," to realize performance near the theorem's limit. These recent developments have been integral in the design of high-speed Internet connections and modern cellular digital telephones.

Example Set #7

An analog television station transmits with a bandwidth 4.5 MHz for the picture. The signal-to-noise power ratio (S/N) at the receiver is typically 100000. The equivalent information rate, R, of an analog video signal is about 4 megabits per second (Mbps), depending on the quality of the signal. How does this compare with the Shannon-Hartley capacity for a television channel? (Recall from rules of logarithms that $log_b(x) = log_{10}(x)/log_{10}(b)$)

$$C = (4.5 \times 10^6)\log_2(1 + 100000) = 75 \times 10^6 \text{ bits per second} = 75 \text{ Mbps}$$

Obviously, the present analog TV system, a standard that dates from 1941, is a very inefficient use of the channel capacity. With a capacity of C = 75 Mbps we get a rate of only about R = 4 Mbps. HDTV systems do little better, achieving a rate of R = 19 Mbps in a 6 Mhz channel with about the same signal-to-noise ratio. The sacrifice of the channel's ultimate capacity is used to reduce the complexity and cost of the television receivers.

A standard land-line telephone circuit has a bandwidth of 3 kHz and a signal-to-nose power ratio (S/N) of 500000. What is the capacity of this communication channel?

$$C = 3000\log_2(1 + 500000) = 57000 \text{ bits per second}$$

Remark: There will be no faster telephone modems than we already have. The V90 standard which allows data rates up to 56000 bits per second was published in 1998. This represents the first consumer electronics device that can achieve data rates near the Shannon-Hartley limit. Although newer modem standards have introduced new features such as fast connection negotiation, the V90 standard will be the last word on modem speed using standard land-line telephone circuits. (DSL service requires special circuits at the telephone company's end of the line to expand the bandwidth of the telephone circuit well beyond 3 kHz.)

12. "1984 WON'T BE LIKE 1984"

Radio broadcasting has proven entertainment value, but efforts to use radio for personal, business, government, diplomatic, and military communications entail a possible loss of privacy. Unless steps are taken to encrypt the communications, anyone with a receiver and access to the signal can listen in. Even wire-line communications can be subject to tapping.

The need to encrypt private messages has a very long history indeed. There is evidence of the use of codes that extends back to the times of Egyptian hieroglyphics. When messages are encrypted then there is also motivation by others to crack the codes and decrypt the message. In the well-known Sunday-school story of Daniel we find that at Belshazzar's feast Daniel was the only one able to read (decode) the handwriting on the wall, "MENE MENE TECKEL UPHARSIN." (Daniel 5:25, KJV) That story is evidence of interest in and awareness of codes even in biblical times.

The role of cryptography in diplomacy and war has a very interesting history.[18] A famous example of codebreaking that is no longer secret happened during World War II when the German Navy used a machine-produced cipher called "Enigma" to communicate with its submarines. The Enigma machine worked much like a typewriter. The plain-text message was typed on an ordinary typewriter-like keyboard and the machine produced the cypher-text. The person receiving the message typed it into another Enigma machine which would then produce the plain-text again. The Allies built another machine they called "Colossus" at Bletchley Park, England, to break the German "Enigma" code. The Colossus machine was essentially a digital computer, although at that time the principles of digital computer were in their infancy. By 1944 the machine was efficient enough to decipher most Enigma coded messages in minutes. As a result the Allies began sinking German submarines at a rate of about one per day.[19][20] The Allies however were not the only ones breaking codes. Through much of World War II the American (in particular) codes used in the field by the Army were weak. After the war it became clear that the Axis had broken into most American field codes, in some cases using machines to decrypt them faster than the intended recipients could. Some believe this cost the Allies thousands of lives at the Battle of the Bulge.

During the Cold War that followed World War II, computer technology developed at an accelerated pace in part supported by research grants from the military, which by then had recognized the strategic importance of superior computer technology. Modern technologies such as today's Internet have strong and easily traced roots in research work funded by the military, especially the U.S. Defense Advanced Research Projects Agency (DARPA). Much of the literature recounting the history of engineering, and especially computing, includes obvious positive attitudes toward a salvific role for computing and technology. In this view, engineering and computing will save us from tyranny, dictatorship, ignorance, tedium, and even from boredom.

An example of hubris in the form of technological positivism is the advertisement that Apple Computer Inc. ran during the 1984 Super Bowl to announce the Apple Macintosh computer. The advertisement opens with a grey and dystopian view of a corporate assembly. At the front of a large auditorium, on a large screen, a videoconference display is running. A black-and-white video image of the leader of the meeting shimmers from the screen. The leader's voice is hollow-sounding in the cavernous auditorium. He says (See Figure 18a), "Today, we celebrate the first glorious anniversary of the Information Purification Directives. We have created, for the first time in all history, a garden of pure ideology where each worker may bloom secure from the pests of contradictory and confusing truths. Our unification of thoughts is more powerful a weapon than any fleet or army on earth. We are one people, with one will, one resolve, one cause. Our enemies shall talk themselves to death and we will bury them with their own confusion. We shall prevail!" Into this totalitarian situation strides a strong young woman carrying a sledgehammer (Figure 18b) that she hurls at the screen in front of the room. Upon impact the screen explodes. As the blast rocks the audience back in their seats a voice-over announcer vociferates, "On January 24th Apple computer will introduce Macintosh. And you'll see why 1984 won't be like '1984.'" (Figure 18c) The second instance of "1984" is a reference to George Orwell's science fiction book, *Nineteen Eighty Four*, first published in 1949 following World War II. The plot of that book involves electroshock therapy and other generally technologic means to pacify and control society.[21] (Some have also associated the leader of the meeting depicted in the advertisement as a representation of the character "Big Brother," in Orwell's novel.)

(a) (b) (c)

Figure 18, Three still shots from Apple Computer's 1984 Super Bowl TV advertisement.[22]

Did the Apple Macintosh computer save the world from a totalitarian future? At least we can say without a doubt that it introduced the "WIMP" (windows, icons, menus, and pointing devices) graphic user interface but that's a rather small accomplishment compared to ending tyranny. Did the Colossus codebreaking machine save the world from dictatorship? It would be hard to deny that is shortened World War II, but dictators and despots still rule in some countries. In each of these examples it is obvious however, that engineers, even whole companies of people, were motivated by their sense of duty to make the world a better place. This inner drive helps engineers and technologists wade through the sometimes-tediously-difficult theory behind technological development. In the end, the product of this technological labor does not save the world from tyranny and the devil—sometimes it even becomes part of the problem. We must not be deluded by the supposed power of technology or conversely, ignorant of our personal responsibility to do technology to the glory of God. Christians must remember that the world belongs to God and we are caretakers, ambassadors (2 Corr 5:20), with finite responsibilities and capabilities. Within that context meaning and purpose can be found.

13. CIRCUIT ANALYSIS

The design of a power system, a cellular telephone, a computer, or any electrical or electronic device relies on an understanding of circuit analysis. This text has already introduced the basic definitions for charge, current, voltage, electrical power, resistance, and voltage sources. With only a little more development some moderately powerful circuit analysis can be done. First, a few more definitions are needed so that we can discuss circuits in a precise way.

Definition: A ***node*** is a model of a portion of a circuit that is assumed to have no resistance. It is the connection between circuit elements.

Definition: A ***bus*** is a collection of nodes that all serve a related purpose. (Occasionally a bus has only one node. This is similar to the idea of a "set" with only one element.)

In a real circuit a wire is often used to make a connection. There are other ways to make a connection, such as a trace on a circuit board or a connection made via a chassis or a metal pipe or a bar of metal, etc. In any case, when we model the circuit, each entire connective part of the circuit is called a *node*. Students sometimes confuse a dot on a schematic with a node. A dot is not a node since it does not symbolize the entire connective area. Each dot is only part of a node. One node can contain several dots. See Figure 19.

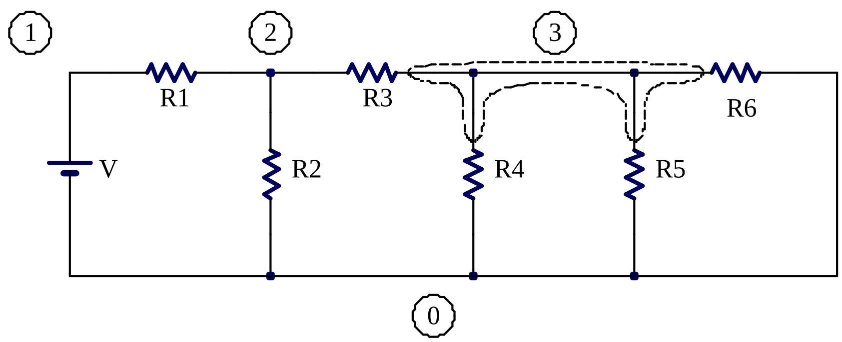


Figure 19. A circuit with four nodes labeled "0" through "3." Node "3" is further illustrated with a dashed line around it. (The six dots in the schematic are not nodes.)

Definition: A ***path*** is a route through a circuit through which current may flow and not enter or exit any node more than once. The direction of travel is considered part of the path.

In Figure 19 one path might be from node "0" up through the battery to node "1" and then to the right through R1 to node "2." Another path simply goes in the other direction, starting at at node "2", going left through R2 to node "1" and then down through the battery to node "0."

A path may not enter (or exit) the same node twice. For example in Figure 19 the route from node "2" right through R3 to node "3" then down through R4 to node "0", then right and up through R5 to node "3" and then right and down through R6 to node "0" is not a path since it enters (and exits) node "3" two times. (It also enters node "0" two times.)

Definition: A ***loop*** (also known as a ***closed path***) is a path that starts and ends at the same node. A loop has a direction.

The direction of a loop is clockwise or counterclockwise in most cases. If the circuit has to be drawn in three dimensions due to crossing but unconnected wires, then simple words like clockwise and counterclockwise might not be adequate to describe the direction of the loop. In that case arrows annotated onto the schematic can be most helpful.

Definition: A ***mesh*** is a loop that does not enclose any circuit elements. A mesh can only be defined on a schematic where no wires cross. (Equivalently, a mesh is a loop that does not enclose any other loops.)

In Figure 19, a path from node "0," clockwise around the outside of the circuit and back to node "0" is a loop. This loop goes through the battery, R1, R3 and R6 (not around them) and then back to the battery. Going over the same route in a counterclockwise direction describes a separate and technically distinct loop. However, just as when we know that $x = 2y$, then stating that $-x = -2y$ is really telling us nothing new—it is the same equation simply multiplied by minus one—so loops that are identical except for direction give us no new information. For this reason, once a loop is defined in one direction, then the corresponding loop in the other direction is ignored when doing circuit analysis, almost as if it did not exist.

The loop that goes clockwise around the outside of the circuit in Figure 19 encloses several circuit elements, namely R2, R4 and R5. Thus it is not a mesh. The loop up through the battery, right through R1, down through R2 and back to the bottom of the battery is a mesh. For each mesh we also choose a direction. The one just described goes in the clockwise direction. As with loops, we ignore the corresponding meshes that simply go in the other direction. Each area of white space on the schematic that is framed by circuit elements gives rise to one useful mesh. The circuit in Figure 19 contains exactly four (useful) meshes. If a schematic is redrawn in a different arrangement that still gives equivalent connections, then a mesh might become an ordinary loop and vice versa. Thus, the concept of a mesh is tightly coupled with the schematic drawing used to define the mesh. If a schematic includes connections that cross over each other, then meshes cannot be defined. Sometimes the schematic can be redrawn with no wires crossing. (In contrast, loops can be defined on any schematic and they do not change if the schematic is redrawn.)

Definition: A ***series connection*** is a connection between *exactly* two circuit elements.

Definition: A ***parallel connection*** is formed when two or more circuit elements share exactly the same node-pair.

In Figure 19 the battery and R1 are "connected in series" because they both connect to node "1" *and there are no other circuit elements connected to node "1."* Note that R1 and R3 are not connected in series because even though they both share a connection to node "2," R2 also connects to node "2." In that same figure, R4, R5, and R6 are "connected in parallel" because they share connections to nodes "3" and "0." R2 and R4 are not in parallel because although they share connections to node "0", they do not share a connection to another node. Note that drawing circuit elements so that the symbols are arranged in a parallel style on the page does not define the type of connection. For example, in Figure 19 R2 and R4 are both drawn vertically—in a parallel style—but they are not connected in parallel. Similarly, R5 and R6 are not drawn in a parallel style, yet they are connected in parallel.

Definition: An ***open circuit*** is a broken, removed, or missing connection.

Definition: A ***closed circuit*** is a connection through which current may flow with practically no resistance. (Also known as a ***short circuit***.)

An open circuit may be deliberate, as when a switch is "opened," or it may be accidental as when something breaks or burns up. Similarly, a closed circuit may be deliberate, as when a switch is "closed." Accidentally closed circuits are usually called "short circuits." Notice that in an electrical circuit an open switch blocks current flow but in a pipe containing a fluid an open valve does just the opposite—it allows current to flow. One must not get confused by the two different meanings of "open." Similarly for "closed."

The above definitions help engineers talk and write more precisely about circuits. They are also helpful in stating two of the most fundamental theorems of circuit analysis, Kirchhoff's laws.

Theorem: ***Kirchhoff's Current Law (KCL)*** states that the algebraic sum of all the currents *exiting* any node is zero.

(An alternative statement of KCL used in some textbooks is that the algebraic sum of all the currents *entering* any node is zero. Yet another version states that the sum of the actual, not algebraic, currents entering a node equals the sum of the actual currents exiting a node. It is the author's opinion that this latter version is especially difficult for students since usually the actual directions of current flow are unknown. Any of these versions works if used properly and consistently, but don't use one version sometimes and the other version other times!)

In the above definition, the word "algebraic" is intended to convey the idea that we will rely only on the direction of the current arrow labels (not the direction of the actual current flow since it may be unknown) to derive a KCL equation. Then a current that actually flows in the opposite direction of the arrow will be accounted for as a negative amount. KCL mentions currents "exiting" a node. Thus, a current actually entering a node will be accounted for as a negative amount exiting the node.

Kirchhoff's current law is based on the principle of conservation of charge. To use a water analogy, consider wires (nodes) to be like pipes. For the analogy to work, we must assume the pipes are filled solid with water. No air may be allowed in the pipes anywhere and the water is assumed incompressible. Maybe the pipes are connected in a "T" shape so there are three openings. Then if 5 gallons per minute of water is flowing out of one opening, we know that the algebraic sum of the flows *out* of the other two openings must be –5 gallons per minute. (The *minus* sign indicates that the net water flow is *not out* of the openings.) The water coming out one opening must be supplied somehow from the other openings because water does not magically appear or disappear inside the pipes. Charge flowing through nodes behaves similarly. It does not magically appear or disappear inside nodes, or circuit elements for that matter. It only flows. As a quantity of charge moves, it pushes the charge in front of it along and drags other charge along behind it, just like water in pipes.

Theorem: **Kirchhoff's Voltage Law** (**KVL**) states that the algebraic sum of all the voltage *drops* around any loop is zero.

(An alternative statement of KVL found in some textbooks is that the algebraic sum of all the voltage *rises* around any loop is zero. Yet another version states that the sum of the actual voltage rises equals the sum of the actual voltage drops around the loop. Again, students tend to have special troubles with this last version since it is difficult to know a priori which voltages are actual rises and which are actual drops. As with KCL, use one version of the theorem consistently in all your work.)

Similarly to KCL, the word "algebraic" here means that we will rely only on the voltage labels (not the actual voltage changes since they may be unknown) to derive a KVL equation. Then a voltage rise will be accounted for as a negative amount of a voltage drop.

Kirchhoff's voltage law is based on conservation of energy. As charge flows through a circuit element, it may do work (a voltage drop) as for a light bulb. Conversely, as charge flows through a circuit element, work might be done on the charge (a voltage rise) as for a battery that is discharging energy into the electrical circuit. Since energy cannot be created or destroyed by current—charge only carries energy—it should be possible to account for all the energy transferred by the charge. If a unit of charge flows one time around a loop, then whatever work it did (voltage drops) must be offset by whatever work was done on the unit of charge (negative voltage drops—voltage rises).

The above Kirchhoff's laws (KCL and KVL) are the keys to traditional circuit analysis. In electrical engineering they have the same fundamental status as Newton's laws (including $F = ma$) do in Statics and Dynamics. (Kirchhoff stated other laws too, such as his laws regarding the radiation of light and heat from hot substances. Additionally, KCL and KVL have several other names such as Kirchhoff's first law, point rule, junction principle, second law, loop law, etc. Although these other names are becoming obsolete, the reader might encounter them, especially in older literature.)

14. SINGLE LOOP CIRCUITS AND KVL

When a circuit contains only one useful loop, then KVL and Ohm's law can be combined to calculate the loop current. From this all voltage drops and the power dissipation of each circuit element can be found. First one labels the circuit with voltage labels. Then a KVL loop equation is written. The variables in the loop equation are replaced as needed by substitutions using Ohm's law so that only one unknown, the loop current, remains in the equation. The equation can then be solved. Then all the other needed variables can be calculated. Here is an example. (See Figure 20.)

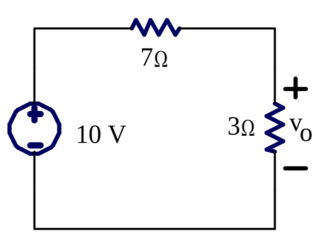


Figure 20. An example of a single-loop circuit. Find voltage v_o.

To apply KVL the figure should be annotated to indicate the direction of the loop and to label any unlabeled voltages. In this case the unknown voltage drop across the 7 Ω resistor needs a label. The annotated figure is shown below. (The annotations are part of the solution of the problem. Problem statements often leave several voltages for you to label.)

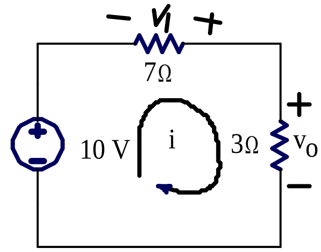


Figure 21. The above figure has been annotated to show loop direction and the unknown voltage v_1.

Now going around the loop in the indicated direction and adding the voltage drops gives:

$$-10 - v_1 + v_o = 0 \qquad \textbf{(Equation 3)}$$

You might wonder how the polarity for v_1 (negative sign on the left) was chosen. It was an arbitrary choice. In this case the voltage across the 7 Ω resistor is unknown. Just put a label down with either polarity, but once labeled, stick with that polarity consistently. Notice the simple convention that the sign in the equation can be found as the first polarity sign encountered when going around the loop.

Notice however that the label as drawn in Figure 21 violates the passive sign convention with respect to the direction of the loop current, *i*. Thus instead of using the equation $v_1 = 7i$ it will be necessary to add a minus sign to correct for the passive sign convention violation. Actually $v_1 = -7i$. In a similar style, $v_o = 3i$ (the passive sign convention is obeyed here). Substituting these values into Equation 3 gives

$$-10 - (-7i) + 3i = 0 \qquad \textbf{(Equation 4)}$$

Solving gives $i = 1$ A. Notice how the two negative signs in front of the "7*i*" cancel each other out in Equation 4. If the v_1 label had been inserted with the other polarity the second "–" sign in the Equation 3 would have been "+" and the sign in the Ohm's law equation would also have been "+". As you can see, either way the answer, $i = 1$ A, is the same. This illustrates that it does not matter which direction you use for the polarity of v_1 so long as you observe the polarity and use the passive sign convention consistently.

After the loop current is found, then the voltage drops can be found. In this case $v_o = 3i$ and we found $i = 1$ A, thus $v_o = 3$ V. Similarly, $v_1 = -7i = -7$ V. (Or, if the polarity had been shown with the plus sign on the left, then v_1 would have been +7 V. Either way, it means the same thing—the terminal on the left has the higher actual voltage.)

The style of this particular loop circuit—two resistors in series connected to a voltage source—occurs rather commonly and therefore has some special significance. If we generalize it we can derive two important concepts. First we observe that "resistors connected in series add." Second, we observe that this circuit "divides the voltage."

Resistors connected in Series Add

The statement that, "resistors connected in series add" is colloquial, but probably a familiar concept from grade school and high-school physics classes. A more precise statement is that, "Resistors connected in series are equivalent to using one resistor which is the sum of all the individual resistors." How can this statement be proven? Consider Figure 22. If R_{eq}, is selected correctly, then given the same voltage source, V, the same current will flow in both circuits. Thus i_{eq} will equal i. This is what we mean by "equivalent."

Definition: Two circuits are **equivalent** if they produce equal currents when the voltages are the same or vice versa, equal voltages when the currents are the same.

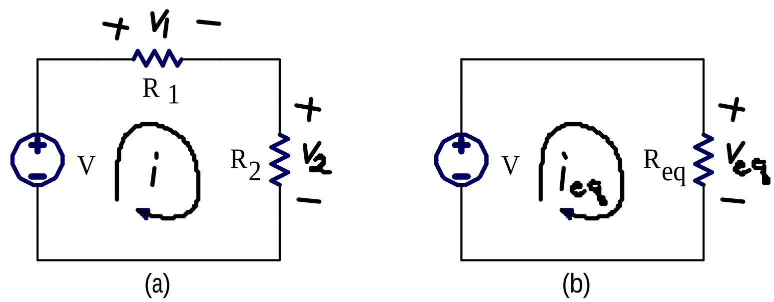


Figure 22. (a) Two resistors in series connected to a voltage source and (b) an equivalent resistance connected to a voltage source.

Writing KVL around each of the two circuits gives:

$$-V + v_1 + v_2 = 0$$
$$-V + v_{eq} = 0$$

Ohm's law gives these equations for v_1, v_2, and v_{eq}:

$$v_1 = iR_1 \qquad v_2 = iR_2 \qquad v_{eq} = i_{eq}R_{eq}$$

Substituting them into the KVL equations gives:

$$-V + iR_1 + iR_2 = 0$$
$$-V + i_{eq}R_{eq} = 0$$

Solving the KVL equations for the current in each case gives:

$$i = \frac{V}{R_1 + R_2} \qquad \textbf{(Equation 5)}$$
$$i_{eq} = \frac{V}{R_{eq}}$$

In these two equations the voltage source is assumed to be the same. Since R_{eq} is supposed to function equivalently to the series combination of R_1 and R_2 we must also let $i_{eq} = i$. Thus, the two KVL equations should be set equal giving:

$$\frac{V}{R_1 + R_2} = \frac{V}{R_{eq}}$$

Solving for R_{eq} gives the final result, proving that "series resistances add":

$$R_{eq} = R_1 + R_2 \qquad \textbf{(Equation 6)}$$

Voltage Divider

A less familiar result called the "Voltage Divider Equation" can be derived from Equation 5 above and from Ohm's law. Substitute $i = v_2/R_2$ in Equation 5 and then solve for v_2. The result is:

$$v_2 = V\frac{R_2}{R_1 + R_2} \qquad \textbf{Equation 7}$$

The above equation shows that a fraction of the original voltage appears across R_2. Since this circuit occurs often enough, the voltage divider equation is sometimes a useful shortcut that wraps up the above KVL circuit analysis in one handy formula.

15. SINGLE NODE-PAIR CIRCUITS AND KCL

In what follows the reader will find many parallels to the preceding section. Where the previous discussion was about loops, the following will now be focused on nodes. Where it was about voltages, it will be focused on currents.

When a circuit contains only one pair of nodes, then KCL and Ohm's law can be combined to calculate the node voltage of one node with respect to the other. From this, all currents and the power dissipation of each circuit element can be found. First one labels the circuit with current labels. Then a KCL equation is written. The variables in the equation are replaced as needed by substitutions using Ohm's law so that only one unknown, the node voltage, remains in the equation. The equation can then be solved. Then all the other needed variables can be calculated. Here is an example (See Figure 23.)

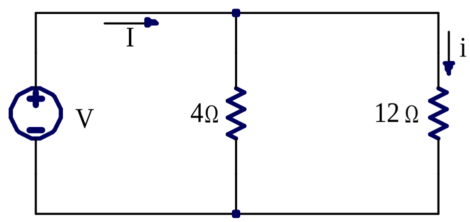


Figure 23. An example of a single node-pair circuit.
Find current, *i* given that the voltage source is adjusted until *I* = 60 A

To apply KCL the figure should be annotated to indicate a reference node (the ground symbol,) and all currents flowing out of the other node toward the reference. In this case it seems natural that the reference node should be the bottom node, however this is actually an arbitrary choice unless some other information is given. In addition, since each resistor is connected by wires to the voltage source, the voltage across each of the two resistors is known to be the same voltage, V, although we do not know exactly how much voltage that is. The annotated figure is shown below. (The annotations are part of the solution of the problem. Problem statements often leave several items for you to label.)

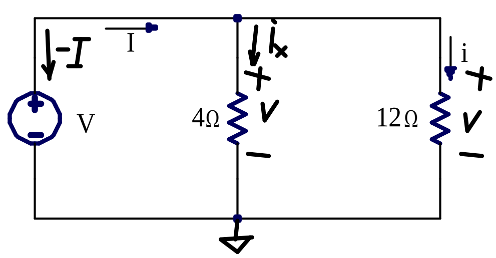


Figure 24. The above figure has been annotated to show a reference node, voltages, and currents exiting the top node.

Now writing a KCL equation at the top node gives:

$$-I + i_x + i = 0 \qquad \textbf{(Equation 8)}$$

You might wonder how the direction for the current arrow label for i_x (downward) was chosen. It was an arbitrary choice. Although perhaps knowing that I = 60 A suggests that i_x might actually flow downward, that cannot be known for sure until the problem is solved. Just put a label down with either direction, but once labeled, stick with that direction consistently.

Now observe that $-I = -60$ A. Furthermore, from Ohm's law observe that $i_x = V/4$ and $i = V/12$. Substituting these values into Equation 8 gives:

$$-60 + \frac{V}{4} + \frac{V}{12} = 0 \qquad \textbf{Equation 9}$$

Solving gives $V = 180$ V. Now that the node voltage has been found all the circuit element currents can be found. In particular, $i = V/12 = 180\text{ V}/12\,\Omega = 15$ A.

The style of this particular node-pair circuit—two resistors in parallel with a known current dividing between them—occurs rather commonly and therefore has some special significance. If we generalize it we can derive two important concepts. First we observe that "the resistance of resistors connected in parallel is equivalent to product over sum." Second, we observe this circuit "divides the current" between the resistors.

Resistors connected in Parallel are Equivalent to Product over Sum

The statement that, "resistors in parallel are equivalent to product over sum" is colloquial, but might be a familiar concept from grade school and high-school physics classes. A more precise statement is that "Resistors connected in parallel are equivalent to the reciprocal of the sum of the reciprocals of all the individual resistances." How can this statement be proven? Consider Figure 25. If R_{eq}, is selected correctly, then given the same current, I, in both circuits, v_{eq} will equal v. This is what we mean by one resistor being "equivalent" to two in parallel.

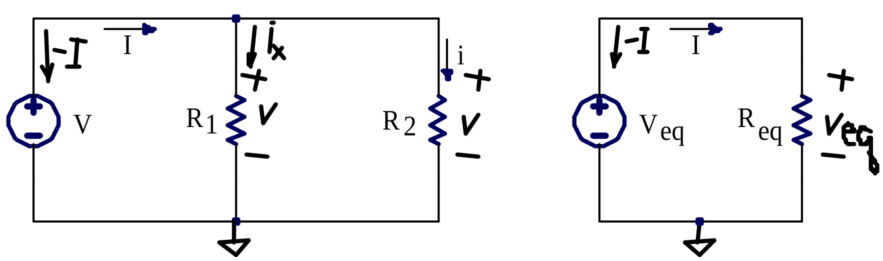


Figure 25. (a) Two resistors in parallel connected to a voltage source (b) an equivalent resistance connected to a voltage source.

Writing KCL at the top node of each of the two circuits gives:

$$-I + i_x + i = 0$$
$$-I + I = 0$$

Ohm's law gives these equations for i_x, i_2, and I

$$i_x = V/R_1 \qquad i = V/R_2 \qquad I = V_{eq}/R_{eq}$$

Substituting them into the KCL equations gives:

$$-I + \frac{V}{R_1} + \frac{V}{R_2} = 0$$

$$-I + \frac{V_{eq}}{R_{eq}} = 0$$

Solving the KVL equations for the voltage in each case gives:

$$V = I\frac{R_1R_2}{R_1 + R_2} \qquad \textbf{(Equation 10)}$$

$$V_{eq} = IR_{eq}$$

In these two equations the current, I, is assumed to be the same. Since R_{eq} is supposed to function equivalently to the parallel combination of R_1 and R_2 we must also let $V_{eq} = V$. Thus, the two KCL equations should be set equal giving:

$$I\frac{R_1R_2}{R_1 + R_2} = IR_{eq} \qquad \textbf{(Equation 11)}$$

Solving for R_{eq} gives the result, proving that "resistors in parallel are equivalent to product over sum":

$$R_{eq} = \frac{R_1R_2}{R_1 + R_2} \qquad \textbf{(Equation 12)}$$

Caution: The above formula only works for *two* resistors in parallel. If there are three or more resistors in parallel a more general formula is needed. It can be found by solving Equation 11 for $1/R_{eq}$ instead of for R_{eq}. This gives:

$$\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2}$$

This equation can be stated as, "Resistors in parallel are equivalent to the reciprocal of the sum of the reciprocals of all the individual resistances." This formula can be extended to any number of resistors in parallel, say n of them.

$$\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \ldots \quad + \frac{1}{R_n} \qquad \textbf{(Equation 13)}$$

Equation 13 can be proved in a style similar to above, just add more resistors in parallel to the first circuit and solve for $1/R_{eq}$ instead of R_{eq}.

Current Divider

A less familiar result called the "Current Divider Equation" can be derived from Equation 10 and from Ohm's law. Substitute $V = iR_2$ in Equation 10 and then solve for i. The result is:

$$i = I\frac{R_1}{R_1 + R_2} \qquad \textbf{(Equation 14)}$$

The above equation shows that a fraction of the total current flows through R_2. Since this circuit occurs often enough, the current divider equation is sometimes a useful shortcut that wraps up the above KCL circuit analysis in one handy formula. Notice that in the current divider formula R_1 is in the numerator (the resistor the current $\underline{i}$ does *not* flow through) whereas in the voltage divider formula R_2 is in the numerator (the resistance the voltage appears across).

16. MESH ANALYSIS

Most circuits are not as simple as having only a single loop or a single node-pair. In these cases there is a variety of more powerful circuit analysis techniques to choose from. Only one, called *mesh analysis*, will be illustrated for this introduction. Mesh analysis, based on KVL, is one of the most popular methods of doing circuit analysis since voltages are usually easy to measure. Although mesh analysis can handle a large variety of circuits, it cannot be used in every possible case. We will leave the other more general methods for later courses.

Consider the circuit shown in Figure 26. It has four nodes and two meshes.

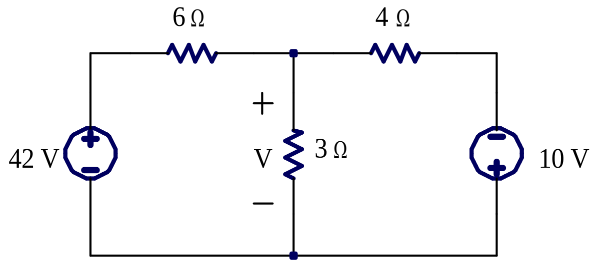


Figure 26. A circuit with four nodes and two meshes. Find *V*.

To solve this circuit by mesh analysis we begin by labeling each mesh and each voltage drop that is not already labeled. The direction of the mesh currents and the voltage polarities usually need to be chosen arbitrarily since we do not know actual directions of current flow and actual voltage polarities until we have solved the circuit. Just as before, if we guess wrong with the labels it will not matter. Double negatives will cancel out and in every case, the meaning of the answers will be the same. A labeled schematic is shown in Figure 27.

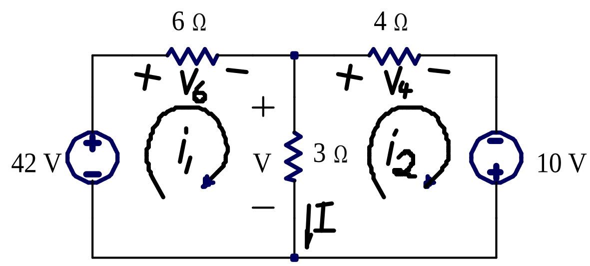


Figure 27. The two-mesh circuit has been annotated to show mesh directions and all voltage drops. A current, *I*, has also been labeled

Now we can write a KVL equation around each mesh.

$$-42 + V_6 + V = 0$$
$$-V + V_4 - 10 = 0$$

Use Ohm's law to rewrite unknown voltages in terms of the mesh currents. for example, $V_6 = 6i_1$ and $V_4 = 4i_2$. The voltage V which appears across the 3 Ω resistor is a more complicated case however. The current flowing down from the top node to the bottom node through the 3 Ω resistor has two components. Current i_1 is flowing down through the 3 Ω resistor and contributes positively, but current i_2 is flowing up through the 3 Ω resistor and thus contributes negatively. It is necessary to distinguish between the element current and the mesh currents.

Definition: The total current that flows in a given direction through a circuit element is called the ***element current*** (Also known as branch current).

Definition: A ***mesh current*** is one component of an element current.

There are at most two mesh currents that compose an element current. This happens every time a circuit element is situated between two meshes. The element current I flowing down through 3 Ω resistor in Figure 27 is an example. Here $I = i_1 - i_2$ where i_1 and i_2 are the mesh currents that compose the element current I. When a circuit element is situated on the outside edge of the schematic, then there is only one mesh current that composes the element current. Mesh currents are an abstraction—that is, you cannot go into a lab, build the circuit shown in Figure 27 and then somehow measure the individual mesh currents flowing through the 3 Ω resistor. You can only measure the element current. Putting all this information together, we see that $V = 3I$, but writing that in terms of mesh currents gives $V = 3(i_1 - i_2)$. Substituting these relationships from Ohm's law into the two KVL mesh equations gives:

$$-42 + 6i_1 + 3(i_1 - i_2) = 0$$
$$-3(i_1 - i_2) + 4i_2 - 10 = 0$$

Further manipulation to collect variables and put the equations into a standard form gives:

$$9i_1 - 3i_2 = 42$$
$$-3i_1 + 7i_2 = 10$$

Now we have two equations in two unknowns. There are several ways to solve such systems of equations. Here we will use a technique called "Gauss elimination." Dividing the top equation by three gives:

$$3i_1 - i_2 = 14$$
$$-3i_1 + 7i_2 = 10$$

Notice that the top equation was manipulated so that the coefficient on the first variable has equal magnitude and opposite sign to the next equation. Then since the left and right sides of the top equation are equal, they can be added to the bottom equation giving:

$$3i_1 - i_2 = 14$$
$$6i_2 = 24$$

Now it is easy to solve the bottom equation. Mesh current $i_2 = 4$ A. That can be substituted back into the top equation which can then be solved for i_1 giving $i_1 = 6$ A. This allows us to find the element current $I = i_1 - i_2 = 6 - 4 = 2$ A. Finally, voltage V can be found as $V = 3I = (3)(2) = 6$ V.

Usually the mesh currents are the unknowns in the KVL equations and the voltage sources and resistances are given, but this is not always the case. Sometimes the goal of an analysis is to solve for the value of a voltage source or resistance such that an element current is a specified amount. Then the mesh equations will include unknown voltages and/or resistances but some of the mesh currents will be known or one mesh current can be written in terms of another and an element current. The goal of the analysis in these cases should again be to find *n* equations in *n* unknowns, where *n* is the number of meshes. Then solve the set of equations by techniques similar to those shown above or other techniques.

Mesh analysis can be extended to handle any number of meshes. Each mesh introduces one equation and one unknown into the analysis. Gauss elimination can also be extended to any number of equations. For example if there were three equations, use the first equation to eliminate the first variable from all the remaining equations. After that, use the second equation to eliminate the second variable from all the remaining equations, etc. Then solve for the last variable, the next to last variable, etc. Once the mesh currents are found, then all the element currents and voltage drops can be found.

Usually the mesh currents are the unknowns in the KVL equations and the voltage sources and resistances are given, but this is not always the case. Sometimes the goal of an analysis is to solve for the value of a voltage or resistance such that an element current is a specified amount. Then the mesh equations will include unknown voltages and/or resistances but some of the mesh currents will be known. (Or one mesh current can be written in terms of another such as $i_1 = i_2 + 10$ A.) The technique to solve these cases is still to find *n* equations in *n* unknowns, where *n* is the number of meshes. Then solve the set of equations by techniques similar to those described above or other techniques.

17. WHAT IS A DIGITAL COMPUTER?

When one appends the adjective, *digital* to the word, *computing*, two kinds of processing are generally expected: 1.) finite arithmetic, and 2.) logical decision-making.

Figure 28. A Chinese abacus. This one is eleven digits wide and can handle up to base-16.[23]

The phrase, *finite arithmetic* implies that there are a variety of kinds of arithmetic, finite arithmetic being one kind of arithmetic. The variety of types of abacuses illustrates that the techniques of doing decimal arithmetic on paper, as taught in grade school, do not span all the systems of doing arithmetic that can be conceived, and that base-two is not the end of the story when it comes to computing numbers. Non-binary computing and *virtualization of machines* and *virtualization of environments* (old, new, or too-complex to be practically built in hardware) is now on the cutting edge of new computer technology. While a "virtual abacus"[24] will probably never be an important computer application, the idea of doing math (or any computing) by techniques that are not presently popular is an important concept that can lead to innovation.

On an abacus (See Figure 28), each rod (column) of beads represents one digit. Each rod usually has two parts, "heaven" and "earth" (or "water") separated by a "beam," although there are some abacuses without a beam. An abacus with eleven columns of beads can represent numbers up to eleven digits wide. An abacus with four earthly beads and one heavenly bead on each rod is designed for decimal arithmetic. It can represent up to ten different digits (0 through 9) on each rod. The number of heavenly and earthly beads determines the base number systems the abacus can work with. A Chinese abacus has two heavenly beads and five earthly beads (or water beads) on each rod. It can work up to base 16 arithmetic, which is handy for dealing with certain systems of measure, such as English pounds and ounces!

The nomenclature of an abacus—heaven, earth, water—reflect ancient systems of cosmology. (Origins and ultimate meaning of life and the creation.) This might strike us as quaintly archaic. However, probably in a future era, the nomenclature of present-day computers—cache, universal serial port, speculative execution, multi-core, pipelining, megaflops, etc.—will be found reflective of our present-era's cosmology. This is an interesting story of our values that will have to be left for other courses such as history and philosophy. The humanities give interesting insights to the ultimate value we ascribe to everyday work! Engineers should look to these non-major courses for insight into engineering itself, not just to round out our personalities.

One often encounters the notion that the word "digital" means the same thing as "binary." Indeed, much digital computing is done these days with binary symbols, but an abacus ought to remind us that this is not necessary. Signal processing using binary symbols has no theoretical advantage over decimal symbols, certainly not if an abacus is used! Digital also has no theoretical advantage over analog! At least in this era, the advantage of digital processing is entirely an economic advantage rooted in the simplicity of the electronic concepts of *on* and *off*, or 1 and 0, or *true* and *false,* as binary values are variously labeled. It is simply less expensive to achieve a particular level of quality using digital signals as opposed to analog signals in most cases.

A modern trend is to use non-binary symbols in communication systems such as cellular telephones, Wi-Fi, Ethernet, HDTV, all of which at present use non-binary signaling. Bases as high as 256 are in practical use right now in situations where the media—optical fiber, copper cable, over-the-air—can economically transport higher-base symbols. (Base 256 implies a set of 256 unique symbols! Usually those symbols are derived from groupings of 8-bits called bytes, {00000000, 00000001, 00000010, 00000011, 00000100, 00000101, . . . 11111111} but each group of eight bits is processed as a single entity, not eight separate bits.) For this course we will focus on binary digital arithmetic, but keep in mind there are other important arithmetic systems, some in other bases. These are in practical use now and there has been such non-binary and non-decimal systems in use throughout history if one looks beyond simple desk-top style computing.

The word, *logic* represents another aspect of what we mean by *computing*. Sometimes decisions need to be made. We desire those to be logical decisions. Theories of logic can be traced back to ancient times in China, India, Greece, and the Islamic world. However in the 1800's George Boole and Augustus De Morgan developed a mathematical process for systematizing logical decisions. George Boole's binary system of propositional logic {true, false} has proven to be especially relevant to modern computing because it coincidentally matches up with the simple electronic concepts of *on* and *off*—another binary pair. However Boolean logic is not the end of the story. There are times when truth values are not known with perfect confidence. One can define logical systems that include uncertainty. One such system is called *fuzzy logic*. Although fuzzy logic can be virtually performed on a modern binary digital computer, it is not native to the computer's design. Fuzzy logic is an important topic in natural speech recognition, artificial vision, fingerprint scanning, and similar technologies. In this course we will focus on Boolean (binary) logic, but keep in mind that there are other important logic systems in practical use. This has also been the case throughout history.

Remember, there is no analog-digital duality in the holistic nature of real life. The main points of this section on what it means to compute is to make as clear as possible to the reader that: a.) Computing does not have to be digital. b.) Digital does not necessarily mean binary. c.) Logical does necessarily mean Boolean logic.

Any fascination with binary digital computing using Boolean logic is a flash of insight (and marketing mania) left over from the past century. Modern designs are increasingly taking advantage of non-digital, non-binary, and non-Boolean designs, even if these designs run as *virtual machines*.[25] To illustrate the value of a virtual machine, a virtual slide rule is just one way that one could conceivably trade away accuracy to gain speed on a binary digital computer. One has to start studying the field of computing somewhere, and binary digital computing based on Boolean logic is a great place to begin, but it is only the beginning.

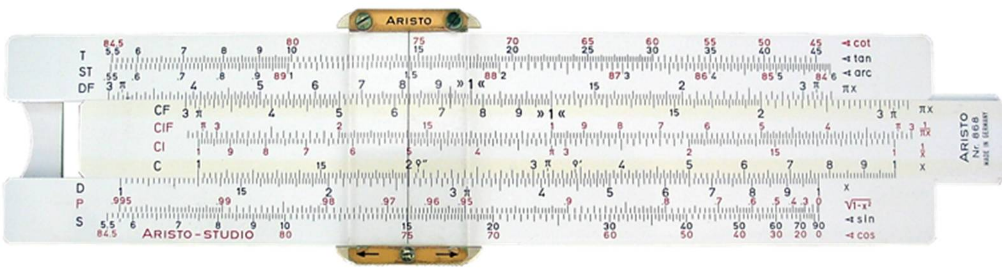


Figure 29. A slide rule. Actual size is about a foot long and 2.5 inches tall. Slide rules are one type of analog computer. This one does multiplication, division, and evaluates trigonometric, exponential, and logarithmic functions accurate to three significant figures.[26]

18. DIGITAL LOGIC USING BINARY SYMBOLS

In the 1850's George Boole outlined a system of algebra that could be combined with propositional statements to make logical decisions. A proposition is a sentence that can be evaluated as true or false. Consider these propositions: *A.) It is raining.* and *B.) The sky is clear.* Suppose that at a particular time on a particular day and at a particular place the truth of each of these propositions is evaluated. First of all, in Boolean logic there are no such things as "drizzling," or "partly cloudy." By the rules that Boole proposed (actually the rules of an algebra), one simply has to settle for the statement being "true," or "false," whatever those might imply. If that seems imprecise, then one needs to refine the propositional statements, and probably use more of them, until one is satisfied. For example, statement *B* could become *B.) Less than 10% of the visible area of the sky is covered by clouds.* Then one could add statement *C.) Less than 90% of the visible area of the sky is covered by clouds.* Now if *B* is false and *C* is true one can conclude that the sky is partly cloudy. (More than 10% and less than 90% covered.) The process of modeling the situation by designing these propositional statements is a process of *abstraction*. Instead of paying attention to every bit of cloudiness in the sky, we reduce our attention to a manageable summary of the available information. We assume that the detail lost by abstraction is not of significance. Thus, we apply our system of values—our worldview—to the situation. It is conceivable that a Christian might do this differently than a non-Christian in some situations, but everyone applies values as they model a situation, whether they do so in conscious awareness or not.

Suppose now that statement *A.) It is raining* is evaluated to be "true," and statement *B.) The sky is clear.* is evaluated to be "false." We can make a binary combination, *A* AND *B* that has the meaning, *It is raining* AND *the sky is clear.* Now to evaluate the truth of this statement Boole proposed that we consider the word AND to be acting as a mathematical operator. Boole defined the operator AND to create a true overall statement if and only if both of the propositions that the operator joins are true. In this example, *A* is "false" and *B* is "true," thus the overall statement, *The sky is clear* AND *it is raining* can be deduced to be *false* by a mathematical calculation. Of course a person can deduce that the statement, *The sky is clear* AND *it is raining* is illogical. Where can the rain come from if there are no clouds? Thus it must be false. But Boole showed how a computer can draw the same conclusion from data provided to it. Boole's system of logic converts propositions and certain words (considered as operators) into a mathematical process that models the way we actually intend the words to function—the way we think logically. We use Boolean logic to enable computers to act, "logically"—by Boole's definition of what is logical.

In this course we will concentrate on just six of the eighteen operators that were important to Boole. Two operators are unary, meaning they operate on only one proposition at a time. The NOT operator is one of those unary operators. For example, if *"It is raining."* is a true statement, then *"It is* NOT *raining"* is a false statement. The NOT operator has operated on exactly one propositional statement and acted to change its truth value. It has become conventional to designate the NOT operator with an overbar. For example, if the Boolean variable A stands for *"It is raining,"* then $\bar{A}$ stands for *It is* NOT *raining.*

Boole also defined an operator that is actually a null operator, or a copy operator. Usually it is called the BUFFER operator. (“Buffering. . .” anyone?) If we “modify” (actually in no way modify, just copy), *It is raining.* to become, *It is raining.* then the truth-value remains unchanged. The definition of the BUFFER operator may seem silly, but it has relevance to making copies of files and signals. The buffer operator is the other unary operator.

A binary operator joins two propositions. Note that this sense of the word “binary” does not have to do with ones and zeros. It has to do with two propositions. The AND operator is a binary operator. How many binary operators might there be? Boole created a table that shows there must be exactly 16 of them. Each logical operator has a name. Some of these names are shown in Table 3. (Trinary and higher operators are reducible to combinations of binary and unary operators, so from a theoretic perspective, they are uninteresting.)

Operator Name ↓	Given $A = 1, B = 1$ Then $Z =$	Given $A = 1, B = 0$ Then $Z =$	Given $A = 0, B = 1$ Then $Z =$	Given $A = 0, B = 0$ Then $Z =$
Contradiction (False)	0	0	0	0
NOR	0	0	0	1
	0	0	1	0
$\overline{A}$	0	0	1	1
	0	1	0	0
$\overline{B}$	0	1	0	1
XOR	0	1	1	0
NAND	0	1	1	1
AND	1	0	0	0
XNOR	1	0	0	1
B	1	0	1	0
	1	0	1	1
A	1	1	0	0
	1	1	0	1
OR	1	1	1	0
Tautology (True)	1	1	1	1

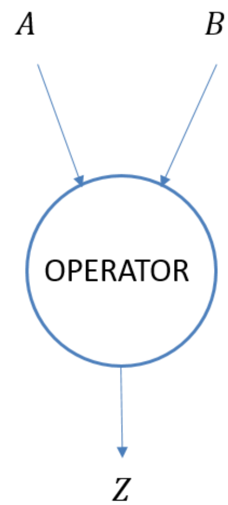


Table 3. A list of all binary Boolean operators.

The binary operators, AND, OR, and NOT, are common. One can show that these three are a sufficient set, meaning that all the other binary operators in Table 3 and BUFFER can be derived from combinations of AND, OR, and NOT. For one example, to get the equivalent of a NAND operation, send the signals *A* and *B* through an AND operation and follow that with a NOT operation. Or, to get the BUFFER operation, send the signal through two NOT operations, one after the other.

Because AND, OR, and NOT are sufficient, in some situations (e.g. a “Boolean search” in a “library database”) these are the only operations supported. However in electronic circuits there are economies to be had by using other logic operations. Other commonly used logic operations are NAND, NOR, and BUFFER. These are the same as AND, OR, and NOT except the logical result is just the opposite. For example, AND is <u>true</u> if and only if the two propositions are true, but NAND is <u>false</u> if and only if the two propositions are true.

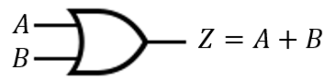
Here is an interesting fact: NAND alone is a sufficient operator. All the other fifteen gates binary gates plus NOT and BUFFER can be made from combinations of NAND gates. NOR is also a sufficient gate. Rather than design three different circuits (AND, OR, and NOT) engineers can design just one circuit, either NAND or NOR, and optimize it to the extreme. Then they can replicate that optimal circuit to create all needed logic operations—millions of replications on one tiny and fast and low-power circuit—to create a CPU.

The function of each Boolean operator can be performed electronically by a circuit called a *logic gate*. The truth value of a proposition is represented by the presence (“true,” or logic-1) or absence (“false,” or logic-0) of voltage. The *logic gate* processes the voltage(s) to perform the desired Boolean operation. A logic gate usually contains just a few transistors. However logic gates can be made in many different ways and do not need to be electrical. There are logic gates that are entirely mechanical (cams driven by a motor as in the pin setting machine at the far end of a bowling lane) or hydraulic (oil in passage ways and pistons as in the automatic transmissions of 1950’s through 1990’s cars) or pneumatic (air pressure in pipes and bellows as in some older thermostats), and in laboratories, there are optical logic gates. Table 4 illustrates the names, symbols, and truth tables of the six most common logic gates.

AND

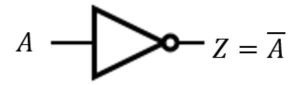
A	B	$Z = AB$
0	0	0
0	1	0
1	0	0
1	1	1

OR



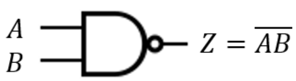
A	B	$Z = A + B$
0	0	0
0	1	1
1	0	1
1	1	1

NOT



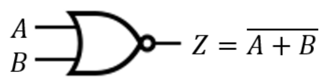
A	$Z = \overline{A}$
0	1
1	0

NAND



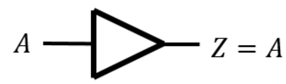
A	B	$Z = \overline{AB}$
0	0	1
0	1	1
1	0	1
1	1	0

NOR



A	B	$Z = \overline{A + B}$
0	0	1
0	1	0
1	0	0
1	1	0

Buffer



A	$Z = \overline{A}$
0	0
1	1

Table 4. The six most commonly used logic gates; their names, symbols, and truth tables.

Example Set #8

Suppose there are three people voting on a policy. The policy will be adopted if a majority of the three people vote “yes.” We desire to design an electronic circuit to compute the vote. (Of course the problem is much more interesting if we have a larger number of votes to process, but we will start out with just three votes!) To apply Boolean logic to this situation we need a suitable set of propositions and we need to figure out the necessary Boolean operations.

The three propositions can be, *A.) Person A votes “yes.” B.) Person B votes “yes.”* and *C.) Person C votes “yes.”* These are legitimate proposition sentences because their truth-value is only “true” or “false.”

A typical error would be to write the three phrases, “*A = person A*, *B = person B*, *C = person C*.” These are wrong because the phrases make no sense. The variable *A* is not in any mathematical way equal to (or not equal to) “person A.” Truth values cannot be ascribed based on the statements. Another possible error would be to write, “The persons are A, B, and C.” This is a proposition—either true or false—but it is irrelevant to the problem. The confusion resulting from improper propositional statements makes it difficult or even impossible to reason out the next step of the solution of the problem.

One way to find a suitable set of Boolean operations is to think about the situation and reduce it to words, then derive the operations from those words. There may be many suitable sets of words, some resulting in more economical circuits than others. There is not necessarily just one “correct answer” to problems like this. (There are better, more complicated methods than illustrated here. Take a course on digital logic to learn more!)

Here we can say that *If A* AND *B* OR *if A* AND *C* OR *if B* AND *C then the majority is “yes.”* Using the mathematical symbol + for the Boolean operation of OR and using the mathematical symbol for implied multiplication as the symbol for the Boolean operation of AND, the truth-value of the vote outcome can be computed as

$$V = AB + AC + BC$$

If you think about it, the above equation says the same thing as the words in the paragraph above. Note that the mathematical symbols do not stand for “equals” or "addition" or "multiplication" in the same sense as they do in high-school math classes. These have been re-defined. However Boolean logic is an "algebra." This means that if you give the symbols for addition and multiplication a more abstract definition, these symbols then share a common definition for both real algebra and Boolean algebra and other algebras. (Take a course on abstract algebra to learn more!) Thus the use of algebraic symbols to represent Boolean logic is not accidental.

The above equation can be reduced to an electronic circuit by connecting gates appropriately. In this case one needs to compute *A* AND *B* (which can be written as *AB*), also *AC*, also *BC*. Then those three computations need to be brought to an OR gate and processed to find the truth-value of *V*. If *V* turns out to be logic-1 or *true* then the policy has a majority vote and should be adopted.

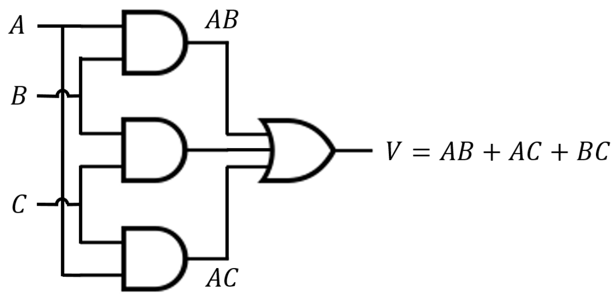


Figure 30. A schematic diagram of an electronic circuit to process the majority vote of three people.

19. COMPUTING WITH NUMBERS

We use symbols for counting. These symbols we call cardinal numbers. (A cardinal number is a symbol or set of symbols used for counting.) Since most people have ten fingers, it has become conventional to count with a set of ten symbols, {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}. We then use a positional placement system to weight the value of these symbols by powers of the base, in this case by powers of ten. This is so familiar to us that we usually give no thought to the underlying mathematical structure of the decimal number system. However, if the base is changed, then the mathematical structure of a positional number system becomes more obvious to our observation. Consider a binary or base-two number system having only the digits {0, 1}. Counting in binary proceeds as follows: 0 (zero), 1 (one)—and then we are out of digits, thus carry to the next place—10 (two), 11 (three)—again out of digits and another carry—100 (four), 101 (five), 110 (six), 111 (seven)—carry again—1000 (eight), 1001 (nine), 1010 (ten) and so forth. The value in decimal notation of the count represented by 101011 in binary can be found from the polynomial expansion in decimal math using the base as follows:

$$1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0$$

$$= 1 \times 32 + 0 \times 16 + 1 \times 8 + 0 \times 4 + 1 \times 2 + 1 \times 1$$

$$= 32 + 0 + 8 + 0 + 2 + 1$$

$$= 43$$

Any integer base can be employed. For the sake of illustration, consider counting in base 17. First of all one needs 17 unique symbols to represent the digits. Let these symbols be the set {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F, G} (Notice that there are seventeen elements in that set.) Counting from zero to thirty-five proceeds as follows: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F, G, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 1A, 1B, 1C, 1D, 1E, 1F, 1G, 20, 21. Notice that since the base is seventeen, 21 does not represent a count of twenty-one. In base seventeen 21 represents a count of thirty-five. This can be seen by using the polynomial expansion:

$$2 \times 17^1 + 1 \times 17^0$$

$$= 2 \times 17 + 1 \times 1$$

$$= 34 + 1 = 35$$

When several bases are in use in nearby contexts it becomes unclear what base is in play. In order to clarify this, the base is identified by postfixing the base as a subscript on the number whenever the base is not otherwise obvious. Thus we know that 21_{10} really does mean twenty-one and that 21_{17} means thirty-five. Sometimes other identifiers are used to indicate the base, such as a postfix of "h" (stands for Hexadecimal) to denote a number as base 16. Another common identifier for base 16 is a prefix of "0x." (That is a zero, then an x.) Still another common identifier for base 16 is % as a prefix. Thus 13h is the same as 0x13 is the same as %13 and all are the same as 13_{16} and by conversion, they are all 19_{10}. An identifier for base-two (binary) that is somewhat common is the prefix of 0b. Thus 0b1011 is the same as 1011_2 and by conversion, the same as all the above, 19_{10}. There are many more such non-subscript methods of identifying the base. In this course we will stick with subscripts.

Any number can be converted to base-ten (decimal) by using the polynomial expansion technique illustrated in the previous paragraphs. However, converting numbers from decimal to some other base requires a different process. Consider converting thirty-five to binary. One can reason in words as follows:

- The largest power of two less than or equal to thirty-five is thirty-two, which is two-to-the-fifth power. Thus a binary 1 is needed in the fifth place, also called the thirty-second's place.
- Subtracting thirty-two from thirty-five leaves three. The largest power of two less than or equal to three is two, or two-to-the first power, thus a one in the first place, or the two's place.
- Subtracting two from three leaves one. The largest power of two less than or equal to one is one, or two-to-the-zeroth power, thus a one in the zeroth place, also called the ones place.

All the other places are filled with a zero. The result is 100011_2 which is equivalent to 35_{10}. All the reasoning above can be done more quickly with a modification of the usual method of doing long division. Read Figure 31 as, "**2 goes into** 35 **seventeen times with a remainder of 1**. Then 2 goes into 17 eight times with a remainder of 1. Etc." When the final quotient is zero, the process is done, although remember to account for the final remainder.

2)35
2)17 R 1
2)8 R 1
2)4 R 0
2)2 R 0
2)1 R 0
0 R 1
100011_2

Figure 31. A modification of long division that will convert base-ten numbers to another base. Conversion to base 2 is illustrated here.

17)35
17)2 R 1
0 R 2
21_{17}

Figure 32. A modification of long division that will convert base-ten numbers to another base. Conversion to base 17 is illustrated here.

When a base greater than ten is used with the above method, one must be careful to insert the correct symbols when needed. Constructing a conversion table as is shown below can be helpful. Consider converting 73683_{10} to base 17. The result, as shown below, is $EGG5_{17}$. (Not 1416165_{17}.)

17)73683
17)4334 R 5 → 5
17)254 R 16 → G
17)14 R 16 → G
0 R 14 E

{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16}
{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F, G}

$EGG5_{17}$

Figure 33. A modification of long division that will convert base-ten numbers to another base. Conversion to base 17 including conversion to proper digits is illustrated here.

Here is an entertaining challenge: Convert 14446013_{10} to base-32. (Hint: Extend the table and method of Figure 33.)

In order to convert from one non-decimal base to another non-decimal base, the most practical method is to first convert to base-ten and then convert to the desired base. For example, to convert 21_{17} to binary, first convert it to decimal. ($2 \times 17^1 + 1 \times 17^0 = 35_{10}$) Then convert to binary as shown in Figure 31. The result is 10011_2. Thus, $21_{17} = 35_{10} = 10011_2$.

The relationship between binary and hexadecimal (base-sixteen) numbers is a special case. The sixteen symbols of the hexadecimal number system, {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F, G} convert to every possible combination of four binary bits. Thus one can convert between hexadecimal and binary by converting each digit individually and then stringing the result together in order of significance. One many find it convenient to add or remove leading zeros during this process. Work from least significant to most significant to reduce confusion as to which binary bits belong in each nibble. (A nibble is a group of four bits.)

Binary	Hexadecimal
0000	0
0001	1
0010	2
0011	3
0100	4
0101	5
0110	6
0111	7
1000	8
1001	9
1010	A
1011	B
1100	C
1101	D
1110	E
1111	F

Start here

Examples: $BEEF_{16} = 1011\ 1110\ 1110\ 1111_2 = 1011111011101111_2$

$EA3_{16} = 1110\ 1010\ 0011_2 = 111010100011_2$

$1000111111010011_2 = 0001\ 0001\ 1111\ 1101\ 0011_2 = 11FD3_{16}$

$10011_2 = 1\ 0011_2 = 0001\ 0011_2 = 13_{16}$

Table 5. Conversions between binary and hexadecimal (base-sixteen)

20. COMPUTING WITH TEXT.

Since computers work on a binary basis, text (letters of the alphabet) has no native representation. In order to do word processing or other text-based tasks letters of the alphabet have to be encoded into bits. There are a number of standard encodings. A popular encoding in the present era is called Unicode. Unicode has been standardized across practically every human language, digital platform, operating system, and computer language. It is a good idea to use Unicode. However because of its broad applicability, it is complicated. In order to illustrate the concept of encoding text, for this course we will use an older but still common encoding called American Standard Code for Information Interchange, or ASCII, which is pronounced like a single word, "ask-ee." For historical reasons, ASCII is a 7-bit code. That is, each symbol (each letter of text, punctuation mark, space, etc.) is encoded as a string of 7 bits. Two to the seventh power is 128. Thus ASCII encodes 128 symbols. An ASCII table is shown in Table 6. The first 32 characters in the ASCII table (codes 0 to 31_{10} or $1F_{16}$) are called *control characters*. Most similar tables, such as Unicode, also include control characters. These are used for file management and machine control, not to convey text. Since most modern computers use word-widths that are a multiple of 8 bits, ASCII codes are usually pre-pended with a leading zero bit to make each ASCII symbol fit into a byte (8-bits). UTF-8 is an ASCII compatible subset of Unicode. Thus this introduction to ASCII is equivalent to learning a small amount about Unicode. A commonly available editor that recognizes ASCII encoding (with some modifications) is Microsoft Notepad.exe, shipped with every edition of Windows. Suppose you type **Hello <Enter> world. Bye.** into a text editor such as Notepad. (<Enter> means striking the "Enter" key. This will case control codes for carriage return—move the cursor to the start of the line—and line feed—move down to the next line—to be inserted into the file.) The editor will store the ASCII byte-codes for each text character one-after-the-other in memory. The codes for this example (in hexadecimal or base-16, in bold red below with their meaning illustrated on the next line down) are:

48	**65**	**6C**	**6C**	**6F**	**0D**	**0A**	**77**	**6F**	**72**	**6C**	**64**	**2E**	**20**	**42**	**79**	**65**	**2E**
H	e	l	l	o	<cr>	<lf>	w	o	r	l	d	.	<sp>	B	y	e	.

If you could place those 18 bytes in a file named with a ".txt" extension, and then open the file in Notepad, you would see what is shown in Figure 34. If you want to try it, take a look at https://hexed.it/ That is a web site that will create a file to your specifications. Size the file to 18 bytes, fill it with the bytes shown above, and export the file back to your computer. Then open it in Notepad.exe. That is exactly how Figure 34 was created.

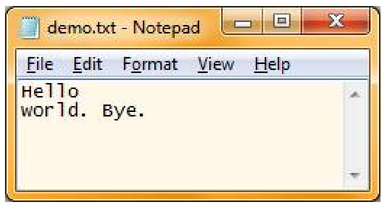


Figure 34. Notepad displaying the file shown above in hexadecimal code.

Decimal	Hex	Char	Decimal	Hex	Char	Decimal	Hex	Char	Decimal	Hex	Char
0	0	[NULL]	32	20	[SPACE]	64	40	@	96	60	`
1	1	[START OF HEADING]	33	21	!	65	41	A	97	61	a
2	2	[START OF TEXT]	34	22	"	66	42	B	98	62	b
3	3	[END OF TEXT]	35	23	#	67	43	C	99	63	c
4	4	[END OF TRANSMISSION]	36	24	$	68	44	D	100	64	d
5	5	[ENQUIRY]	37	25	%	69	45	E	101	65	e
6	6	[ACKNOWLEDGE]	38	26	&	70	46	F	102	66	f
7	7	[BELL]	39	27	'	71	47	G	103	67	g
8	8	[BACKSPACE]	40	28	(	72	48	H	104	68	h
9	9	[HORIZONTAL TAB]	41	29	)	73	49	I	105	69	i
10	A	[LINE FEED]	42	2A	*	74	4A	J	106	6A	j
11	B	[VERTICAL TAB]	43	2B	+	75	4B	K	107	6B	k
12	C	[FORM FEED]	44	2C	,	76	4C	L	108	6C	l
13	D	[CARRIAGE RETURN]	45	2D	-	77	4D	M	109	6D	m
14	E	[SHIFT OUT]	46	2E	.	78	4E	N	110	6E	n
15	F	[SHIFT IN]	47	2F	/	79	4F	O	111	6F	o
16	10	[DATA LINK ESCAPE]	48	30	0	80	50	P	112	70	p
17	11	[DEVICE CONTROL 1]	49	31	1	81	51	Q	113	71	q
18	12	[DEVICE CONTROL 2]	50	32	2	82	52	R	114	72	r
19	13	[DEVICE CONTROL 3]	51	33	3	83	53	S	115	73	s
20	14	[DEVICE CONTROL 4]	52	34	4	84	54	T	116	74	t
21	15	[NEGATIVE ACKNOWLEDGE]	53	35	5	85	55	U	117	75	u
22	16	[SYNCHRONOUS IDLE]	54	36	6	86	56	V	118	76	v
23	17	[ENG OF TRANS. BLOCK]	55	37	7	87	57	W	119	77	w
24	18	[CANCEL]	56	38	8	88	58	X	120	78	x
25	19	[END OF MEDIUM]	57	39	9	89	59	Y	121	79	y
26	1A	[SUBSTITUTE]	58	3A	:	90	5A	Z	122	7A	z
27	1B	[ESCAPE]	59	3B	;	91	5B	[	123	7B	{
28	1C	[FILE SEPARATOR]	60	3C	<	92	5C	\	124	7C	|
29	1D	[GROUP SEPARATOR]	61	3D	=	93	5D	]	125	7D	}
30	1E	[RECORD SEPARATOR]	62	3E	>	94	5E	^	126	7E	~
31	1F	[UNIT SEPARATOR]	63	3F	?	95	5F	_	127	7F	[DEL]

Table 6. ASCII conversion table—to and from decimal or Hex (base-16).

Word processors, spreadsheets, presentation slide editors, e-mail clients, Web browsers, and other programs all use text codes similar in concept to ASCII code, but larger in table size and usually with embellishments such as "escape sequences" and so forth. Their *binary files* (ending in extensions such as docx and pptx, and so forth) are simply strings of bytes such as in the Notepad file illustrated in Figure 34. The programs used to edit these files are designed around the appropriate code table(s). Some file types have two levels of encoding. Hypertext markup language (.htm or .html) is an example. Special codes, written in ASCII text, such as **`<br>`** to start a new line of text, are inserted into an otherwise ordinary text file. (Which is why you can edit

.html files with a text editor if you wish—and if you know the html codes.) For example, type **<html><body>Hello
world. Bye.</body></html>** into a notepad window and save the file with the extension .html. Then open the file in your favorite Web browser. The web browser will read the bytes in the code and convert them to ASCII text, then read the html code from the ASCII text and render it (decode it) in the Web-browser's window. You will see something similar to Figure 35.

By the way, the hexadecimal code for this .html file is

```
3C 68 74 6D 6C 3E 3C 62 6F 64 79 3E 48 65 6C 6C 6F 3C 62 72 3E 77 6F
72 6C 64 2E 20 42 79 65 2E 3C 2F 62 6F 64 79 3E 3C 2F 68 74 6D 6C 3E
```

You could type that into a hex editor such as https://hexed.it/ and save that file with a .html extension, and get the same result when you open it in your Web browser. (Or you could open it from Notepad.exe and see the raw hypertext markup code.)

Figure 35. Chrome displaying the hypertext markup code shown in the paragraph above.

By now one might be wondering how videos (possibly 3-D), photographs, sound clips, line drawings (e.g. computer-aided-drafting), equations with many math symbols (like the integral sign) and other objects are stored and manipulated in a computer. Those are subjects beyond the scope of this introduction to electrical engineering, but those objects are encoded into streams of bytes just as text files are. Also, there is the matter of compressing files so that they can be stored using fewer bytes. Some compression is lossless, which is achieved by translating from one code, say ASCII, to another code that uses fewer bits to represent the information. Many codes for audio, video, and still photos are designed to remove information that humans would not usually perceive. This results in some distortion in the resulting output, but can save even more storage space. All these matters fall in the area of computer engineering and information theory, which are sub-disciplines of electrical engineering. The theories behind these codes and compression methods are a blend of physics (what is color?), biology (how does the human eye work and what can it actually see?), psychology (what is bothersome and what is pleasing?), math (how to quantify and code), and more. An engineer benefits from a broad education in many subjects.

21. CONCLUSION

This introduction has only begun to scratch the surface of the field of electrical engineering, a rich and influential branch of learning. For over a century electrical engineers have been refining the theories and methods of the discipline. The result has enabled profound cultural developments such as reliable electric power for light, heat, cooling, and even safety systems. Radio and Computers have enabled communications that easily span the college campus, the world, and even outer space. Entirely new social relationships have become possible, such as social media provides. The economy is facilitated by electronic auction sites and now the fledgling appearance of electronic currency (e.g. Bitcoin). We have seen that many engineers, electrical engineers included, are highly motivated by a personal desire to make the world a better place, yet we find that this is a serious challenge as every new technology brings with it unintended consequences. A Christian worldview helps keep the challenge, the risks, and the possibilities in perspective. For those called, the field of electrical engineering fulfils a deep-seated sense of Christian purpose—purpose that is found by faith for it is by faith alone that we are justified. Worship is not just for Sunday, it is for everyday engineering too. May the favor of the Lord rest upon us (Psalm 90:17).

REFERENCES

[1] Astronomy Picture Of The Day, http://antwrp.gsfc.nasa.gov/apod/ap001127.html, accessed 1/7/2008.

[2] Elgerd, Olle I., *Electric Energy Systems Theory, An Introduction*, Second Edition, McGraw-Hill Book Company, 1982, ISBN 0-07-019230-8.

[3] Wikimedia Commons, https://commons.wikimedia.org/wiki/File:Galvanometer_scheme.svg and https://commons.wikimedia.org/wiki/File:D%27Arsonval_ammeter_movement.jpg, accessed 1/9/2018.

[4] Bureau International des Poids et Mesures, http://www.bipm.org/en/si/base_units, accessed 1/7/2008.

[5] Wikimedia commons, https://commons.wikimedia.org/wiki/File:Workers_in_the_fuse_factory_Woolwich_Arsenal_Flickr_4615367952_d40a18ec24_o.jpg, accessed 1/9/2018.

[6] Niagara Falls Thunder Alley, http://www.niagarafrontier.com/image/HSTmilldistrict.jpg, accessed 1/7/2008.

[7] Thomas Cole, *A Distant View of Niagara Falls*, oil on panel, 1830, part of the "Friends of American Art" collection at the Art Institute of Chicago. (Photo by Naomi De Boer.)

[8] Niagara Falls Public Library, http://www.nfpl.library.on.ca/nfplindex/show.asp?id=99770&b=1, accessed 1/7/2008.

[9] U.S. Government photo, "Yangtze River, Three Gorges Section," public domain, https://commons.wikimedia.org/wiki/File:YangtzeInThreeGorges.jpg, accessed 1/9/2018.

[10] Georg Ohm, *Die galvanische Kette: mathematisch bearbeitet (The Galvanic Circuit Investigated Mathematically)*, Berlin, Riemann, 1827. Available, http://www.mb.fh-nuernberg.de/bib/textarchiv/Ohm.Die_galvanische_Kette.pdf, accessed 1/7/2008.

[11] The Institute for Electrical and Electronics Engineers, http://www.ieee.org, accessed 1/7/2008

[12] The Institution of Engineering and Technology, http://www.theiet.org, accessed 1/7/2008.

[13] The Wireless Operators, White Star Ships, http://www.titanic-whitestarships.com/Carpathia%20Rescue.htm, accessed 1/9/2018.

[14] The Wireless Operators, White Star Ships, http://www.titanic-whitestarships.com/Carpathia%20Rescue.htm, accessed 1/9/2018.

[15] White, Thomas, "Building the Broadcast Band," http://www.oldradio.com/archives/general/buildbcb.html, accessed 1/7/2008

[16] Mark Durenberger, "Clear Channel Station Articles," originally published in the *Radio World Newspaper*, June 2000 and following issues, available http://www.oldradio.com/archives/general/ accessed 1/7/2008.

[17] FCC Public Notice, "Mass Media Bureau Announces Revised AM Expanded Band and Filing Window for Eligible Stations," March 17, 1997, available http://www.fcc.gov/mb/audio/decdoc/pdf/da97-537.pdf accessed 1/7/2008.

[18] David Kahn, *The Codebreakers*, The Macmillan Company, New York, 1967. Available from the John and Louise Hulst Library, Dordt College.

[19] Public Broadcasting System, *Decoding Nazi Secrets*, originally broadcast November 7, 1999. Available from the John and Louise Hulst Library, Dordt College.

[20] Wladyslaw Kozaczuk, *Enigma: How the German Machine Cipher was Broken, and How it was Read by the Allies in World War Two*, University Publications of America, 1984. Available from the John and Louise Hulst Library, Dordt College.

[21] George Orwell, *Nineteen Eighty Four*, Harcourt Brace, 1949. See also http://en.wikipedia.org/wiki/Nineteen_Eighty-Four

[22] Apple Computer Inc., Television Commercial. Available http://en.wikipedia.org/wiki/1984_(television _commercial), accessed 1/7/2008.

[23] Abacus, Pixabay.com, https://pixabay.com/en/abacus-count-mathematics-485705/ accessed 1/9/2018.

[24] Acula Online Calcuators, Online Abacuses, http://www.alcula.com/calculators/abacus/ accessed 1/11/2018.

[25] For an entertaining recreation on non-binary computing read Issac Asimov's short story, "The Feeling of Power." (1958, February) Worlds of IF (Science Fiction) Magazine, pages 4-11, 115, https://archive.org/stream/1958-02_IF#page/n5/mode/2up accessed 1/9/2018

[26] Slide Rule, Wikimedia Commons, https://en.wikipedia.org/wiki/File:Sliderule_2005.png accessed 1/9/2018.