Introduction to Engineering for Electrical and Computer Engineering Students

Dr. Richard R. Johnston, Professor of Electrical and Computer Engineering,
Lawrence Technological University

Table of Contents

To the Student

The purpose of this course is to introduce you to some of the concepts and issues with which you will deal in your study of, and subsequent practice of, electrical and computer engineering. There was a time when virtually every engineering freshman had considerable experience with electronic kits, or amateur radio, or had disassembled an automobile engine and successfully reassembled it. As our technology becomes more and more complicated, there are fewer opportunities to learn about technology by assembling and disassembling common objects. This course attempts to introduce you to concepts that previous generations of engineering students, living in a time of less complicated technology, learned on their own before arriving at the university.

To the Instructor

The sections on Word, Mathcad, and Maple are included for reference only. It is expected that you will open live Word, Mathcad, and Maple documents when you discuss these topics.

There is a rich selection of material included here, which is a little ambitious for a 2 credit course. Some picking and choosing among the topics and activities is usually necessary. You should be continually aware of this so that you can make decisions about what to keep and what to eliminate as early as possible.

Charge, Current, Voltage, Energy, and Power

Physicists describe five fundamental forces (four, depending upon how one counts them) in nature. They are, in order of familiarity in everyday experience: the gravitational force, the magnetic force, the electric force (charge), the weak nuclear force (or simply the weak force), and the strong nuclear force (or simply the strong force). Modern physicists typically list only four forces, counting the electric force and the magnetic force as manifestations of the single underlying electro-magnetic force. However at this stage, in keeping with the original, separate, experimental description of these forces we shall list five.

Gravitation

We are familiar with the gravitational force, quantified by Newton as,

$$f = \frac{G \cdot m_1 \cdot m_2}{s^2}$$

where f is the force, m_1 and m_2 are the masses involved, and s is the distance between the center of mass of the two masses involved in the gravitation. The unit of force in SI (mks) units is the Newton, in U.S. common units as pound-force or simply, pound. Whenever the unit “pounds” is used alone it ALWAYS means pounds-force. The meaning of the above equation is that we hypothesize the existence of a characteristic of objects we call “mass” whose result is that objects having mass are attracted to one another by the force given by Newton’s gravitation equation. The unit of mass is the kilogram in SI units, and the slug or pounds-mass in U.S. customary units, where 1 slug = 32.2 pounds-mass. The choice of mass units affects the numerical value of the constant G in Newton’s gravitation equation.

Energy (symbolized by W) is the work done applying a force over a distance, s.

$$W = \oint \vec{f} \cdot d\,\vec{s}$$

which reduces to

$$W = f \cdot s$$

when the motion is collinear to the force. The unit of energy is the N-m. The – is a hyphen, not a minus sign. We read it as the Newton meter, understanding a multiplication of Newtons times meters. The N-m is also known as (a.k.a.) the Joule.

Force and motion are related by Newton’s famous second law of motion

$$f = m \cdot a$$

where a is the acceleration, i.e. the change in velocity. Note that the application of a force results not in a velocity (a speed) but a <u>change</u> in velocity. Velocity is given as distance divided by time (distance per unit time)

$$v = \frac{s}{t}$$

and has SI units meters per second, and U.S. common units of feet per second or miles per hour. Acceleration is given as velocity divided by time (velocity per unit time)

$$a = \frac{v}{t} = \frac{\left(\frac{s}{t}\right)}{t} = \frac{s}{t^2}$$

The unit of acceleration is, therefore, (in SI units) meters per seconds squared.

It is worth noting at this point that

$$\frac{\left(\frac{a}{b}\right)}{c} \neq \frac{a}{\left(\frac{b}{c}\right)}$$

since

$$\frac{\left(\frac{a}{b}\right)}{c} = \frac{a}{b} \cdot \frac{1}{c} = \frac{a}{b \cdot c}$$

and

$$\frac{a}{\left(\frac{b}{c}\right)} = \frac{a}{1} \cdot \frac{c}{b} = \frac{a \cdot c}{b}$$

By convention a/b/c (without parentheses) is taken to mean

$$\frac{\frac{a}{b}}{c} = \frac{\left(\frac{a}{b}\right)}{c}$$

From Newton's second law we have

$$a = \frac{f}{m_{Object}}$$

and from Newton's law of universal gravitation, we have (for objects at sea level)

$$f = \frac{G \cdot m_{Object} \cdot m_{Earth}}{{radius_{Earth}}^2}$$

so objects near the surface of the earth are accelerated toward the center of mass of the earth as

$$a = \frac{G \cdot m_{Earth}}{radius_{Earth}{}^2} = 9.8 \cdot \frac{m}{s^2} = 32.2 \frac{ft}{s^2} = g$$

So an object at a distance s above the earth (where s << $radius_{Earth}$) experiences a force f,

$$f = m \cdot g$$

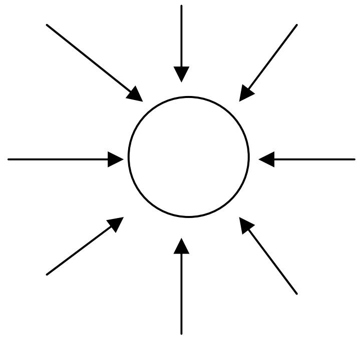
where f is independent of s and g is a constant. We are all familiar with the fact that if an object is raised above the ground and released, that the object falls to the ground. If the object has mass m, and is raised a distance s above the ground then work is done by gravity is

$$W = f \cdot s = m \cdot g \cdot s$$

While the object is sitting at height s, it has the potential to do work if it were released. We say that the object has potential energy U,

$$U = m \cdot g \cdot s$$

We ordinarily abstract the notion that objects released near the surface of the earth fall to earth by saying that the earth somehow modifies the space near the earth such that objects in that space fall toward the earth. We say that the earth imposes a “gravitational field” on the space around it, which we usually picture (in the simpler two-dimensional case) as



The interpretation of this picture is that a mass released at some position would fall toward the earth along the line connecting the position from which it is released to the center of the earth. We can characterize two positions having different distances from the surface of the earth as possessing a difference in potential energy. (The potential energy U is higher for the position that is further from the earth and lower for the position that is closer to the earth.) Notice that this difference in potential (energy) is strictly a function of position, and that zero potential energy is DEFINED to be that of an object at the surface of the earth.

Charge

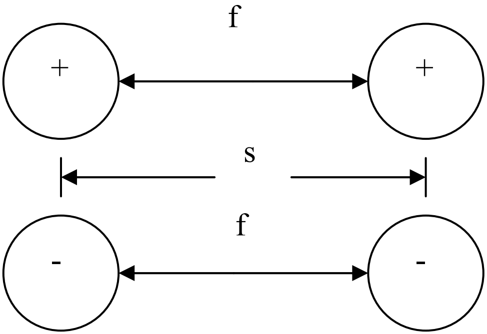
The next most familiar force for most of us is the magnetic force, but because it is not central to our initial discussion of electricity, and because the magnetic force is somewhat more difficult to quantify than the electric force, we shall delay our study of magnetism to a later chapter and consider the source of the electric force; charge.

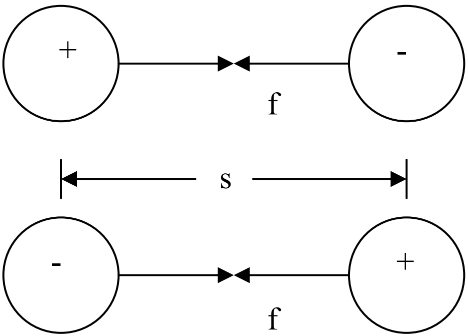
Most of us are familiar with the experimental results that require us to hypothesize the existence of charge. The ancient Greeks knew that if they rubbed a piece of amber in a certain way, that it would attract small bits of hay or papyrus. Most of us have seen someone rub a balloon on their hair and stick it to a wall. The central result that makes the electric force different from the gravitational force is that the force is sometimes attractive and sometimes repulsive. Therefore, we need a characteristic that can come in two flavors: so that we may explain why the force is sometimes attractive and sometimes repulsive. We call the characteristic of matter that produces the electric force "charge," and say that it can be either positive or negative. If we have two objects separated by a distance s (center-of charge to center of charge), they experience a force given by Coulomb's equation:

$$f = \frac{q_1 \cdot q_2}{4 \cdot \pi \cdot \varepsilon_0 \cdot s^2}$$

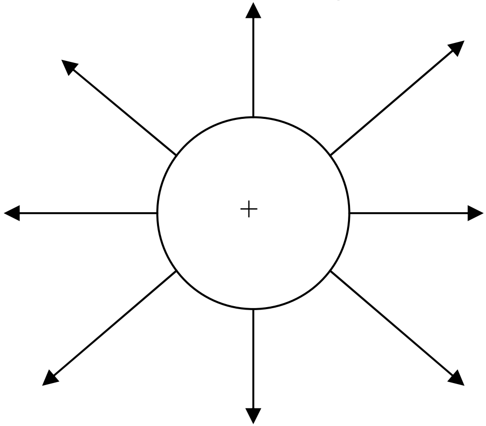
where q_1 and q_2 are the charges in Coulombs, and ε_0 is the permittivity of free space.

Note that since the product of two negative numbers is positive (as is, of course, the product of two positive numbers), and the product of one positive number and one negative number is negative, we may interpret positive values of f given by Coulomb's equation as repulsive, and negative values of f as attractive as shown below.


f
+
+
s
f
-
-


+
-
f
s
-
+
f

We envision the existence of a charged object as modifying the space around it by imposing an electric field as shown below. (Again, this is a 2-D simplification.)



We can characterize two positions having different distances from the surface of the charge as possessing a difference in potential energy. (The potential energy V is lower for the position that is further from the charge and higher for the position that is closer to the charge.) Notice that this difference in potential (energy) is strictly a function of position. (Note the change in the point of zero potential energy in the repulsive charge force case from the one in the attractive gravitational force case. This location of zero potential energy is a matter of definition and is usually chosen for convenience; in cases like the above, zero potential is usually defined at infinite distance.)

The electric potential energy difference has units Joules per Coulomb, a.k.a. volts.

(After Alexander Volta who constructed the first battery.) This difference in potential energy (usually referred to as, simply, difference in potential) between two points in space is one of the fundamental quantities in electrical and computer engineering.

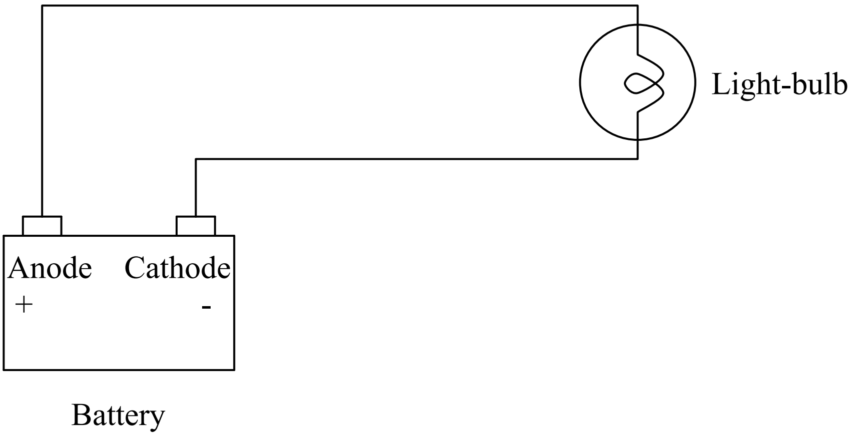
Current

If we have two uniform charge distributions, one positive and one negative, in proximity as shown below, then an electric field is set up as is also shown below.

The interpretation of the above figure is that if a positive charge is let go in the field it would travel to the right. If we count the amount of charge, in Coulombs, moving to the right per unit time (second), we can define the current, I where I has units Coulombs per second, a.k.a. Amperes. (After Andre Ampere who quantified electric current.) We shall not undertake a study of the nuclear forces here, except to point out that the strong nuclear force is what holds protons together in the nucleus (Remember that the protons are all positively charged, so that enormous charge forces must be overcome to keep the protons together in the nucleus.), and that the weak nuclear force governs some kinds of radio-active beta decay.

Batteries and Light-bulbs (lamps)

Now all of this preliminary work is necessary to explain the results of one of the central electrical engineering experiments that we shall undertake. Suppose we take a battery and connect a light-bulb to it using copper wire as shown below.



What we observe, of course, is that the light-bulb gives off energy in the form of light and heat. This energy must come from the battery, and our task is to quantify this phenomenon sufficiently to allow us to make predictions as to how much energy is transferred from the battery to the light-bulb. To explain this transfer of energy, we say that the battery voltage causes a current flow in the wire, and that this current flow is the mechanism by which energy is transferred from the battery to the light-bulb. (We'll have more to say about battery chemistry, and how it causes voltages later.) Before we investigate the current in the copper wire at an atomic level, we pause to consider the units of voltage and current in terms of fundamental quantities.

$$Volts = \frac{Joules}{Coulomb} \qquad Amperes = \frac{Coulombs}{\sec ond}$$

$$V \cdot I = Volts \cdot Amperes = \frac{Joules}{second}$$

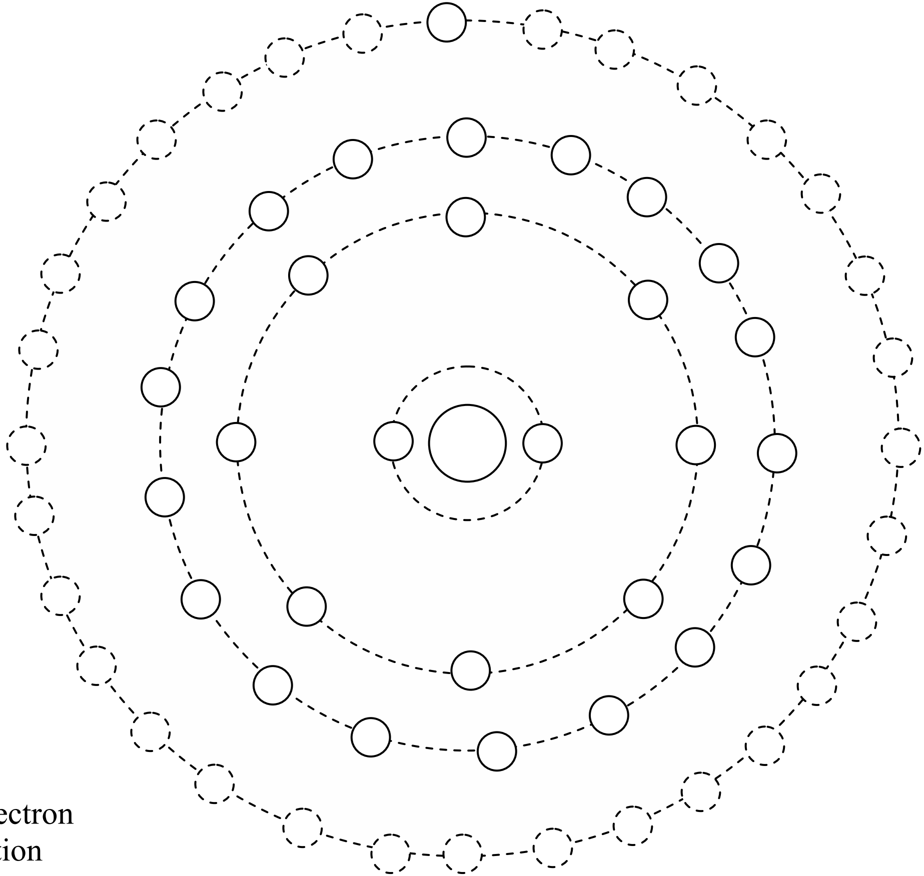
Power, the time-rate-of-change of energy, has units Joules per second a.k.a Watts. (After James Watt inventor of the steam engine.) We may then write our first fundamental Electrical Engineering formula (which you must know at least as well as your own name), the **fundamental power formula**:

$$P = V \cdot I$$

Now we have considered current flow in a simplified 2-D universe which contained only two charge distributions (to produce an electric field) and a few charges which could move under the influence of that field. The reality of current flow in a copper (or aluminum) conductor is considerably more complicated. In fact, it is so complicated that an tractable picture of current flow at the atomic level is well-nigh impossible. It would be like trying to picture the trajectories of individual gas molecules whose random movement is responsible for the phenomenon we call pressure. Or imagine that a wire is like a long narrow cage crowded full of thousands of hyperactive gerbils running furiously in every direction. We could cause "gerbil current" to flow by throwing gerbils in one end, and pulling them out of the other end. We would have a net flow of gerbils, but no guarantee that any gerbil we threw in one end would ever make it to the other end. We shall investigate the flow of current in copper, since it is by far the most common material used for making electrical conductors. Copper has atomic number 29 and atomic weight 63.6. It consists of a nucleus of 29 protons and either 34 or 36 neutrons and an electron cloud of 29 electrons. Copper has two stable isotopes: Cu 63, which contains 34 neutrons, and Cu 65, which contains 36 neutrons. Naturally occurring copper is about 69.17% Cu 63 and 30.83% Cu 65. The copper electrons have different energies associated with them, as well as, different affinities for their (the electrons') nucleus. In the simplified Bohr (After Niels Bohr who suggested it) model of the atom, we imagine that the electrons are little spheres orbiting the nucleus, and associate different orbits to the different energy levels that are predicted by quantum theory. The electrons are arranged in the following way.

Energy Level N	Shell	Sub-shell	Possible Electrons	Actual Number of Electrons
1 (2 Electrons)	K	s	2	2
2 (8 Electrons)	L	s	2	2
		p	6 6	
3 (18 Electrons)	M	s	2	2
		p	6 6	
		d	10	10
4 (32 Electrons)	N	s	2	1
		p	6 0	
		d	10 0	
		f	14	0

Notice that there is only one electron in the N shell, and that its orbit is far from the nucleus. Quantum theory predicts that the atom's affinity for complete energy levels is even stronger than the atom's affinity for charge neutrality. The N-shell electron is therefore only loosely bound to its nucleus, and moves easily under the influence of an externally imposed electric field. When an external electric field is imposed across a conductor (by connecting it to a battery, for example), the loosely bound electrons are free to move and thereby produce an electric current. But, we must not imagine that there is a smooth flow of electrons in the wire like marbles flowing through a hose. When an electron is moved by the external electric field, it cannot travel very far (not more than a few angstroms (1 angstrom – 10^{-10} m), before it collides with another copper atom, where it may or may not be captured depending upon its velocity and original energy.



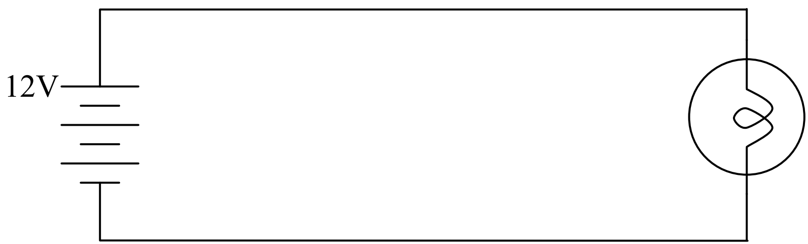
Copper Electron
Configuration

In addition, the source of the external electric field must supply electrons to the conductor at the negative end and remove electrons at the positive end to keep the current flowing. If one calculates the net (average) speed of the electrons flowing in a conductor one finds that they are traveling at about 3 miles per hour: the average walking speed of a human being! Since these motions are so complicated and can only be described statistically we prefer to assume that there is a smooth flow which we shall call current. One of the first stumbling blocks most students encounter is the notion of conventional current. Conventional current (the smooth flow that we are defining) consists of a smooth flow of POSITIVE charges, and therefore flows from positive to negative. Since the actual charge carriers are negative, many students are annoyed at the convention chosen for the nature of conventional current. The usual explanation given in introductory courses, that Benjamin Franklin simply guessed incorrectly when he assumed that electric current consists of positive charge carriers – so now we're stuck with it, isn't the whole story – or even the main part of the story. Conventional current is chosen so as to produce equations that have as few negative signs in them as possible. The rationale is that every negative sign in an equation is another chance to make a mistake. Also, secondarily, the equations look better (are more beautiful) without all the extra negative signs. For almost all our circuits (interconnected collections of components) we are going to assume that the conductors are perfect conductors having no resistance to the flow of current (more about the idea of resistance presently) and are therefore equi-potential surfaces. That is, there is no difference in electric potential energy between the two ends of our conductors.

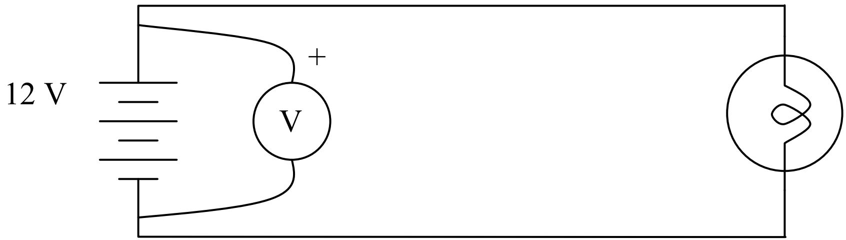
Now voltage (difference in electric potential energy) and current (flow of charges) are three-dimensional phenomena. We are going to simplify things by making a two

dimensional simplification called the lumped element model. Our circuits consist of collections of two terminal elements (some elements have more than two terminals, but we delay study of these until later) connected together by ideal (perfect) conductors. We only define electric potentials at the terminals of the devices (elements): we restrict the number of points in 2-D space where potentials (and therefore differences in potential – voltages) are defined. Further, currents are constrained to flow only within the devices. In order for this to work, it is necessary that the only electric fields in the circuit are those produced by one or more of the elements.
Let us return to our light-bulb circuit. We draw what are called schematic diagrams (or simply schematics) using special symbols for the various elements as shown below.

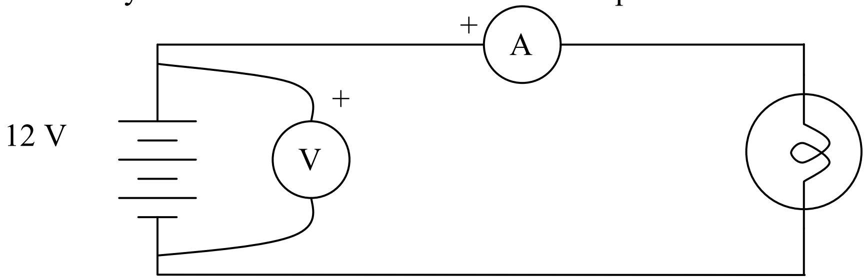
Circuits



We would like to know the electric power being absorbed by the light-bulb from the battery. In order to calculate the power of the light-bulb P, we need to know the voltage across the light-bulb V, and the current through the light-bulb I. Once we have these two numbers we may characterize the light-bulb. To find the information we need, we introduce two measuring instruments: the ideal voltmeter (for measuring voltages), and the ideal ammeter (for measuring currents). Now a voltage exits across an element, so the voltmeter is placed across the element to measure voltages as shown below.



However, an ammeter measures currents which flow THROUGH elements. To use an ammeter you MUST break the circuit! Correct placement of the ammeter is shown below.



NEVER connect an ammeter across a device like a voltmeter. If you do, you will DESTROY the ammeter. We'll have more to say about why this is momentarily.

Note the plus signs associated with the meters: they designate the positive terminal of the meter, which is usually colored red (the other terminal of the meter is usually colored black).
Notice that because the wires are equi-potential surfaces, the voltage across the light-bulb is the same as the voltage across the battery. Also the light-bulb current is constrained by the connections to be the same as the current coming out of the battery. (Note that there only two different electric potentials in this circuit, one at the top where the positive terminal of the battery is connected to the top of the light-bulb, and the other at the bottom where the negative terminal of the battery is connected to the bottom of the light-bulb.)
The power in the light-bulb is then

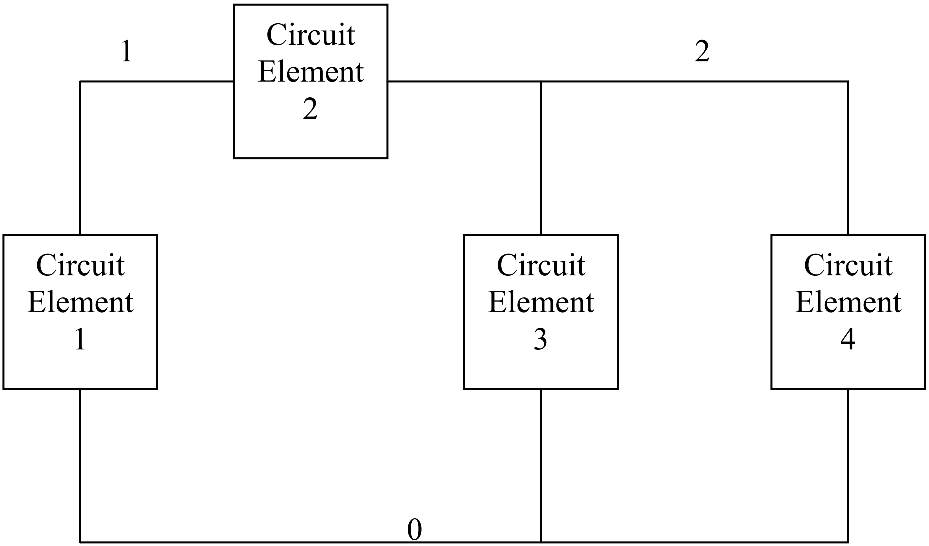
$$P = V \cdot I$$

It is conventional to describe the way the light-bulb influences the current by defining its resistance R. Resistance (to the flow of electric current) is defined by **Ohm's law** (After George Simon Ohm) given below in the same form as Ohm published it.

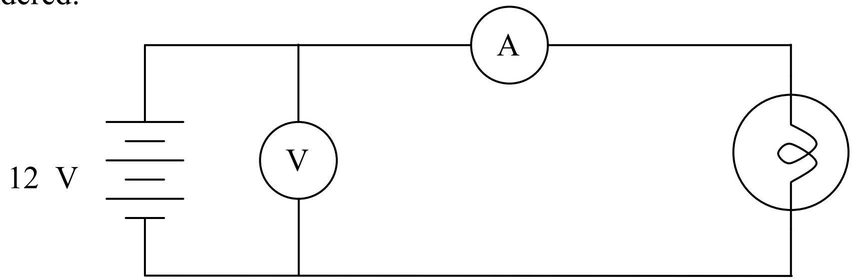
$$R = \frac{V}{I}$$

In most texts Ohm's law is usually stated in a slight algebraic rearrangement as:
$V = R \cdot I$ This fundamental relationship MUST be committed to memory. The unit of resistance is the Ohm, symbolized by Ω. Note that both Ohm's law and the fundamental power formula require that the current enter the terminal that has the higher (more positive) electric potential. In this first circuit, the positive terminal of the battery is

clearly more positive (by 12 volts) than the negative terminal of the battery so our use of Ohm's law and the power formula for the light bulb are justified. In order to solidify our notions of potential differences and currents (and to settle the matter of the location of zero potential), consider the circuit below.



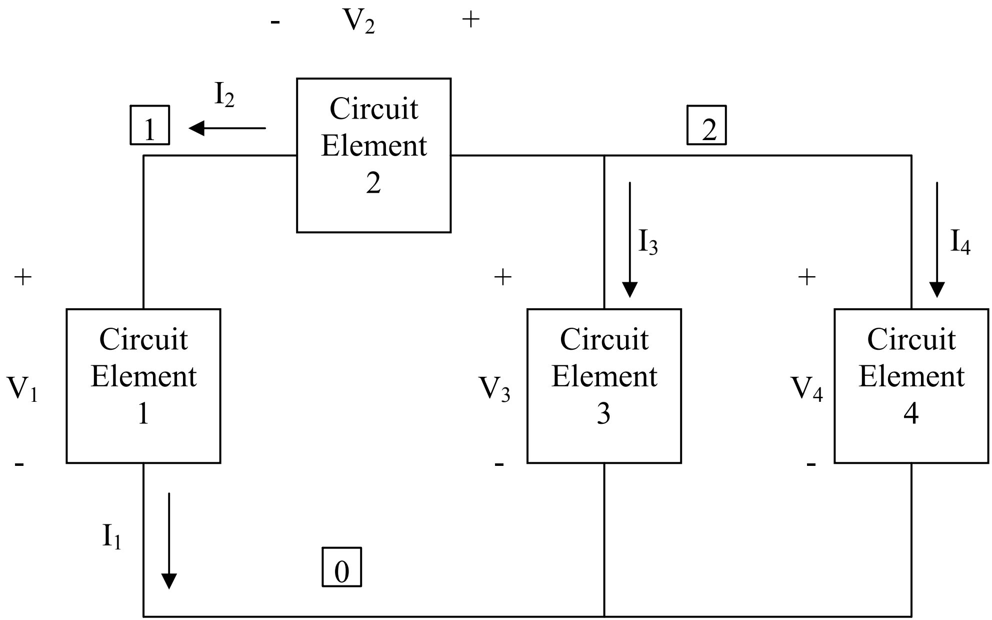
Notice that there are only three distinct electric potentials in this circuit, and they exist at the points labeled 0, 1, and 2. These locations where distinct potentials may exist are called the nodes of the circuit. We call any point where two or more circuit elements are connected together a node. We are free to choose any node and define zero electric potential there. The node so chosen is called the reference or the ground node. (The ground potential is defined as zero volts.) By convention the bottom node is usually chosen and labeled 0. In a standard schematic drawing all circuit elements are either vertical or horizontal, never diagonal. Also connecting wires are always either vertical or horizontal, never diagonal or curved. Occasionally in a textbook unusual configurations (like the ones we used in showing the proper placement of the meters) are used to call attention to special details. However when you draw a schematic (whether by hand or by computer) you should follow the rules as given. The ammeter schematic would then be rendered:



The earlier drawing emphasizes the location where the voltmeter is to be connected. Notice that the circuit connections impose certain constraints on the voltages and the currents that are possible in the circuit. For example the current in element 1 must be the same as the current in element 2 by the conservation of charge. Two elements that have the same current (sometimes we say "share the same current") are said to be in **series**. Notice that the difference in potential across element 3 must be the same as the difference in potential across element 4 (the wires are equi-potential surfaces). Elements that have the same voltage (difference in potential) are said to be in **parallel**. Sometimes we say that parallel elements share the same voltage.

Kirchoff's Circuit Laws

Let us define element voltages and currents as shown below.

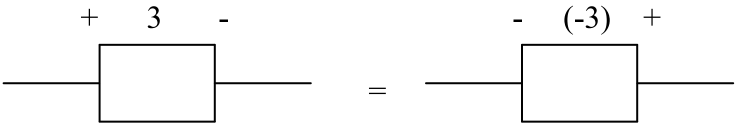


Note that conservation of charge requires not only that $I_1 = I_2$, but $I_2 + I_3 + I_4 = 0$. Hence **Kirchoff's current law**: the sum of the currents entering a node (in this case zero) must equal the sum of the currents leaving the node. In the course in circuits you will be shown a more useful, compact way of stating and using Kirchoff's current law.

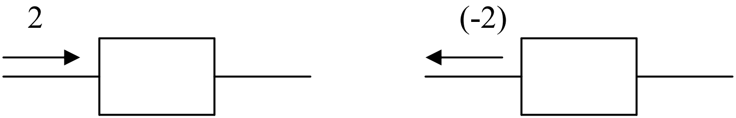
Suppose that we were able to "jump into" the schematic and sit at the bottom of element 1 at node 0. Let us suppose further that we embark on the following journey. We shall travel from node zero (the bottom of element 1) to the top of element 1 (node 1), from there to the right side of element 2 (node 2), and finally back to the bottom of element 3 (node 0). Let us notice the electric potential at various points in our journey. When we travel from node 0 to node 1 the electric potential energy increases (becomes more positive, becomes less negative – these terms mean the same thing) by amount $|V_1|$. (We are traveling from negative to positive.) When we travel from node 1 to node 2 the electric potential energy again increases but this time by an amount $|V_2|$. Finally when we travel from node 2 to node 0 the electric potential energy decreases by $|V_3|$. However we are now back at zero electric potential energy, so we must have

$$V_1 + V_2 = V_3$$

We state **Kirchoff's voltage law**: the sum of the voltages around any closed loop is zero. We point out that a difference in potential can be stated in more that one way as shown below.



Currents can also be described in two ways, as shown below.



Ideal Meters

Now that we have characterized voltage, current, power, and resistance to electrical current, we can return to the characterization of the ideal voltmeter and ideal ammeter. An ideal meter should never change the circuit into which it is inserted, otherwise our reading would not reflect the circuit we are trying to measure, but the circuit as altered by the presence of the meter. The voltmeter is placed across devices and must do its job without drawing any current from the circuit. We, therefore, define the ideal voltmeter as having infinite resistance: its current is, therefore, zero. The ammeter is placed in series with elements and must do its job without altering any of the voltages in the circuit. We, therefore, define the ideal ammeter as having zero resistance: its voltage is, therefore, zero.

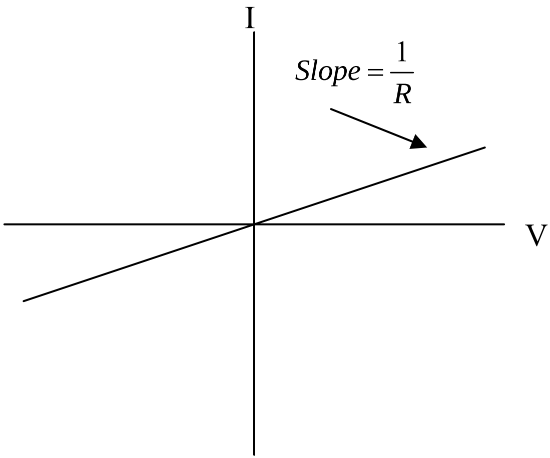
Ideal Resistors

We have already said that the light-bulb presents a resistance R to the flow of current given by

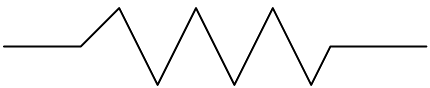
$$R = \frac{V}{I}$$

Now it is perfectly possible (and it is true for the light-bulb) that the quantity $\frac{V}{I}$ gives a different result for each different value of V. We say that the light-bulb does not obey Ohm's law because the ratio $\frac{V}{I}$ is not constant $\forall V$. (The mathematical symbol $\forall$ is read "for all.") This is true because Ohm assumed that the resistance R is a constant, not a function of V. If a device that does not obey Ohm's law is operated under constraints that keep the voltage within a fairly narrow range, then, under those conditions, Ohm's law can still give approximate results that are useful, if not completely general.

The device that DOES obey Ohm's law is called the ideal resistor. If the current through an ideal resistor is plotted for various voltages across the resistor, then the I-V curve shown below is generated. We say that the resistor is characterized, or defined by its I-V curve.

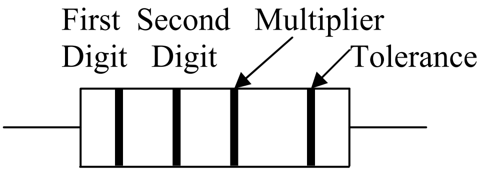


The curve is drawn with V on the horizontal axis to emphasize the causality: a voltage is placed across a resistor – current, therefore, flows. The resistor has the following schematic symbol:



A properly drawn resistor has three "humps."

Resistors come in standard values which are indicated by colored bands. The two most popular standard series are the 5% resistors and the 1% resistors. (Resistors formerly came in 10% series and even 20% series.) The manufacturer guarantees that the actual value of a 5% resistor will be within 5% of its nominal value. For example a 560 Ω, 5% resistor could have a resistance between 560 – 0.05(560) = 532 Ω and 560 + 0.05(560) = 592 Ω. A 560 Ω 1% resistor would be guaranteed to have a resistance between 554.4 Ω and 565.6 Ω. Of course, manufacturers do not set out to make 5% or 1% resistors, the set out to make 560 Ω resistors. After manufacture the resistors are tested, and the ones within 1% go in the 1% bin, and the ones within 5% go in the 5% bin, and the rest are either discarded or perhaps sold as 10% resistors (if they are within 10% of the nominal value). A 5% resistor has four colored bands that are interpreted as follows.



Color	Digit	Multiplier	Tolerance
Black	0	1 (10^0)	
Brown	1	10 (10^1)	1%
Red	2	1,000 (10^2)	2%
Orange	3	10,000 (10^3)	3%
Yellow	4	100,000 (10^4)	4%
Green	5	1000,000 (10^5)	
Blue	6	10^6	
Violet	7	10^7	
Grey	8	-	
White	9	-	
Gold		0.1	5%
Silver		0.01	10%
No Band			20%

So, for example, a color sequence of green, blue, brown, gold, would be a 560 Ω**,** 5% tolerance resistor. A 1% resistor has five bands which will be discussed in a subsequent course. Notice that three of the bands are together at one end of the resistor; these are the value bands, and are read from the one closest to the end of the resistor.
Resistors come in standard values, which for 5% resistors are:
10 11 12 13 15 16 18 20 22 24 27 30 33 36 39 43 47 51 56 62 68 75 82 91

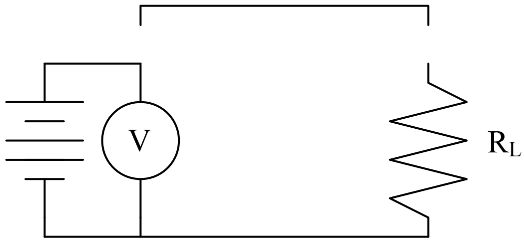
This means that0.22; 2.2; 22; 220; 2200; 22,000; 220,000; etc. are all standard values: the exact value depends on the color of the multiplier band. Electronic components are usually described by units using the SI prefixes. 22,000 would be 22k, 220,000 would be 220k, 2,000,000 would be 2M. The prefixes and their meanings are given in the table below.

Prefix	Name	Meaning
f	femto	10^{-15}
p	pico	10^{-12}
n	nano	10^{-9}
μ	micro	10^{-6}
m	milli	10^{-3}
k	kilo	10^{3}
M	Mega	10^{6}
G	Giga	10^{9}
T	Terra	10^{12}

You will find that the physical size of a resistor is unrelated to its resistance: large value resistors are not physically larger than small value resistors. The physical size of a resistor is related to another specification of the resistor: namely the resistor's power handling capacity. The power rating is the amount of power the manufacturer guarantees that the resistor can safely dissipate. The standard values are ¼ W, 1/3 W, ½ W, 1W, and 2 W. Larger power rated resistors are referred to as power resistors.

Thevenin's Theorem

When we solved the battery and light-bulb problem, we assumed that the battery's voltage did not change when we connected it to the battery; i.e., if the battery voltage was 12 V before we connected it to the light-bulb, we assumed that the battery voltage was 12 V after we connected it to the light-bulb. In other words we considered the battery as a constant voltage source. We shall define an **ideal voltage source** as a source whose voltage stays constant no matter how much current we draw from it. Such an element is very useful, since it simplifies so many calculations considerably. Unfortunately such a device does not exist. We can come pretty close, but all real voltage sources, whether batteries or electronic supplies, have terminal voltages that decrease when current is drawn from them. Thevenin was pondering this problem, trying to find a way to continue to use the notion of an ideal voltage source, while still accurately describing the voltage decline when current is drawn from the voltage source. His elegant solution was to imagine a resistance in series with the voltage source. He called this imaginary resistance the internal resistance of the battery. Notice that this "resistance" is purely fictitious – if you cut open a battery you will NOT find a resistor inside! The internal resistance is a simple way to account numerically for the effects that the battery chemistry has on the operation of our circuit. The internal resistance of a battery can be found by the following experiment. Build the following circuit, but do NOT connect the wire from the battery to the load yet, and measure the battery voltage – call this value $V_{Unloaded}$. The resistor symbol is conventionally used to represent circuit loads.



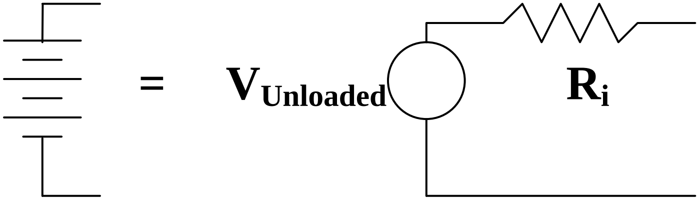
Now connect the load and measure the battery voltage again – call this value V_{Loaded}. The battery's internal resistance R_i can be found from the formula

$$\frac{R_L}{R_L + R_i} \cdot V_{Unloaded} = V_{Loaded}$$

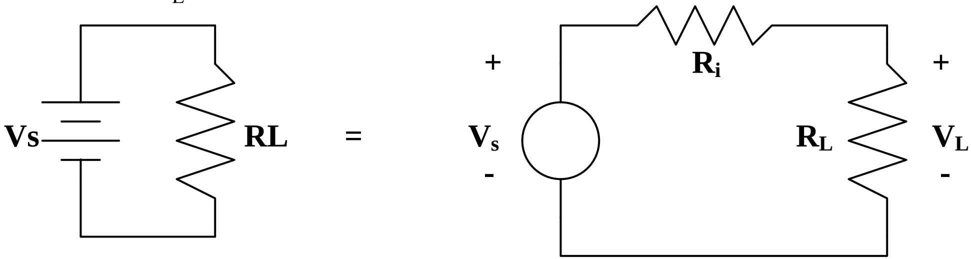
You will derive this formula in your first circuits course. (It depends on the voltage divider equation, which you also will study later.) A little algebraic manipulation isolates R_i as

$$R_i = \frac{R_L}{V_{Loaded}} \cdot \left(V_{Unloaded} - V_{Loaded} \right)$$

We have, then, the following circuit model of the battery.



If we have a battery, whose internal resistance we know, connected to a load, whose resistance we know, then we can find the voltage at the load by the voltage divider equation (which should be committed to memory) as follows. For simplicity and for the sake of generality, we shall name the unloaded battery voltage V_s, and the voltage across the load as V_L.

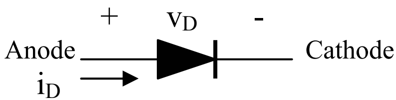


$$V_L = \frac{R_L}{R_i + R_L} \cdot V_S$$

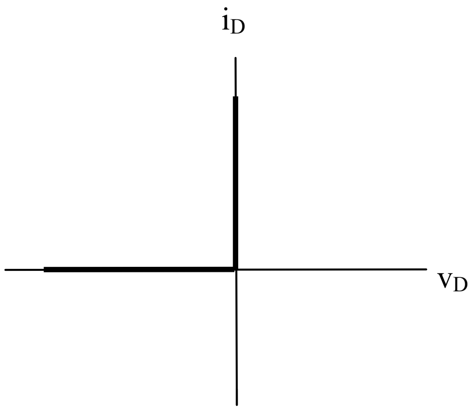
This equation is valid in general for any series combination of a voltage source (known voltage) and two resistors.

Diodes and Light-emitting Diodes (LEDs)

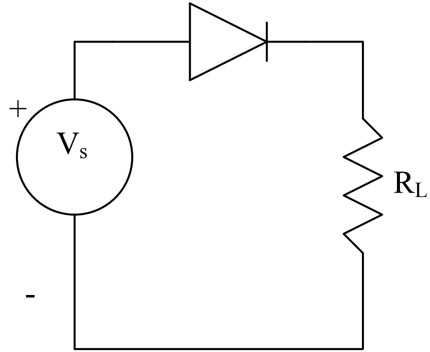
There is a circuit element, called a diode, that permits current flow in one direction but not the other. (The LED is a special sort of diode that emits light whenever current flows through it.) The diode has the following schematic symbol:



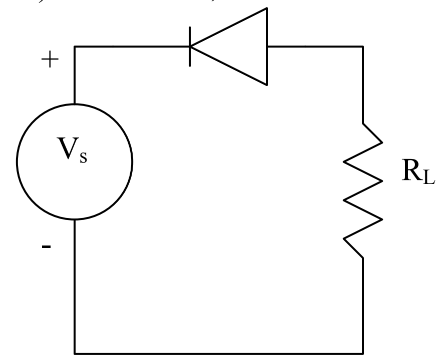
Current flows only in the direction of the arrow, from anode to cathode. The **ideal diode** is characterized by the following I-V curve.



The graph is interpreted as follows. Whenever the diode voltage is more positive on the anode and more negative on the cathode, the diode is said to be forward biased and whatever current the rest of the circuit flows while the voltage across the diode remains zero. So if we had



We say that the diode is forward-biased (the positive terminal of the battery is connected to the anode (the positive terminal) of the diode, and therefore current = V_s / R_L flows. If, on the other hand, we had

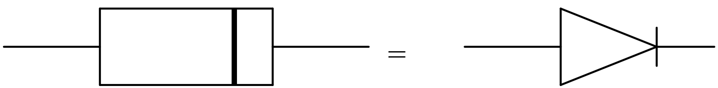


We say that the diode is reverse biased (the positive terminal of the battery is connected to the cathode (the negative terminal of the diode), and no current flows.

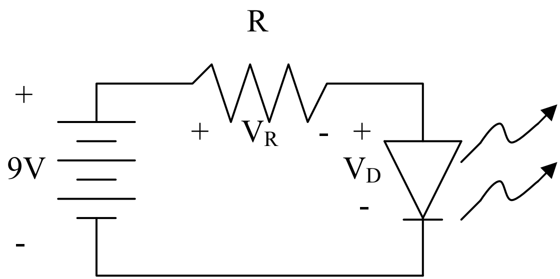
Real Diodes

Unfortunately we cannot build an ideal diode; a real diode approximates the behavior of the ideal diode with the following limitations. A real diode under sufficiently strong forward bias has a small, relatively constant (largely independent of variations in the diode current) voltage across it. This voltage can vary from 0.3 V to 2 or 3 V, depending on the diode, and the amount of current flowing through the diode. A reverse biased diode conducts a very small current (the reverse saturation current) that is largely independent of variations in the diode voltage, for a sufficiently strong reverse bias. This current varies from a few pico-amperes to a few hundred nano-amperes, depending on the diode and on the strength of the reverse bias. Diodes have many uses in electronics: steering currents, rectification (transforming AC to DC), detectors in radio circuits, logic devices, circuit protection, etc., etc.

Ordinary diodes come in packages similar to resistors with a band at one end. The end with the band is the cathode.



At this stage in our study of electricity we are interested in diodes that emit light whenever current flows through them. Such diodes are called light-emitting diodes (LEDs). These diodes are widely used as pilot lamps, and (as we shall use them) as voltage indicators. We would like to power our LED with a 9 V battery, but most LEDs require that the voltage across the diode be 2 V, and that the current be 20 to 60 milli-amperes(mA). The light output of an LED depends directly on the power being absorbed by the LED – more power means more light. If we connected the diode directly across the 9 V battery, not only would the diode voltage be wrong, but the diode would draw a very large current from the battery destroying the diode. We solve both these problems by placing a resistor in series between the battery and the diode as shown below. A resistor used in this way is usually referred to as a current-limiting resistor. (The squiggly arrows differentiate an LED from an ordinary (non light-emitting) diode



The question is "How do we choose the value of the resistor?" The answer lies in the specifications of the LED. Let us suppose that the manufacturer's specification for the

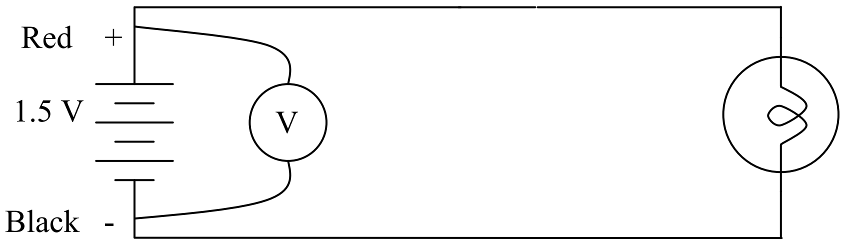
LED calls for 2.0 V across the diode with a diode current of 50 mA. If we have 2.0 volts across the LED, then KVL (Kirchoff's Current Law) requires that the voltage across the resistor must be 9 – 2 = 7 V. Since the LED and the resistor are in series, the LED current and the resistor current must be the same. So we have a resistor whose voltage should be 7 V and whose current should be 0.05 Amperes (50 mA). The resistor's value is, by Ohm's law:

$$R = \frac{V}{I} = \frac{7}{0.05} = 140\,\Omega$$

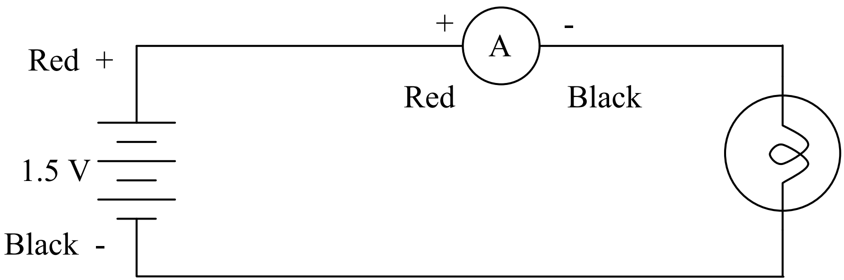
The two closest standard values (the only 5% values that we can buy) are 130 Ω and 150 Ω . We ordinarily choose the resistor that results in a current less than the specification. For R = 150 Ω, I = 46.7 mA.

Laboratory Exercises

1. Construct the following circuit and record the difference in potential (voltage) across the light-bulb.



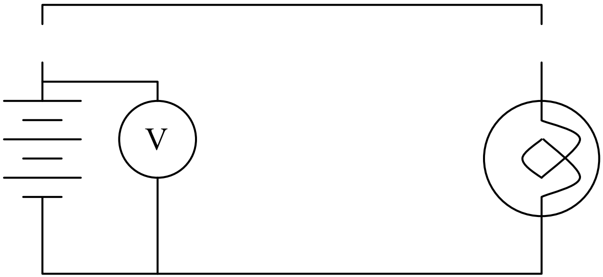
2. Construct the following circuit and record the current flowing through the light-bulb.



Calculate the hot resistance of the light-bulb and the power delivered to the light-bulb.

3. Repeat the above two exercises (1. and 2.) using the motor in place of the light bulb.

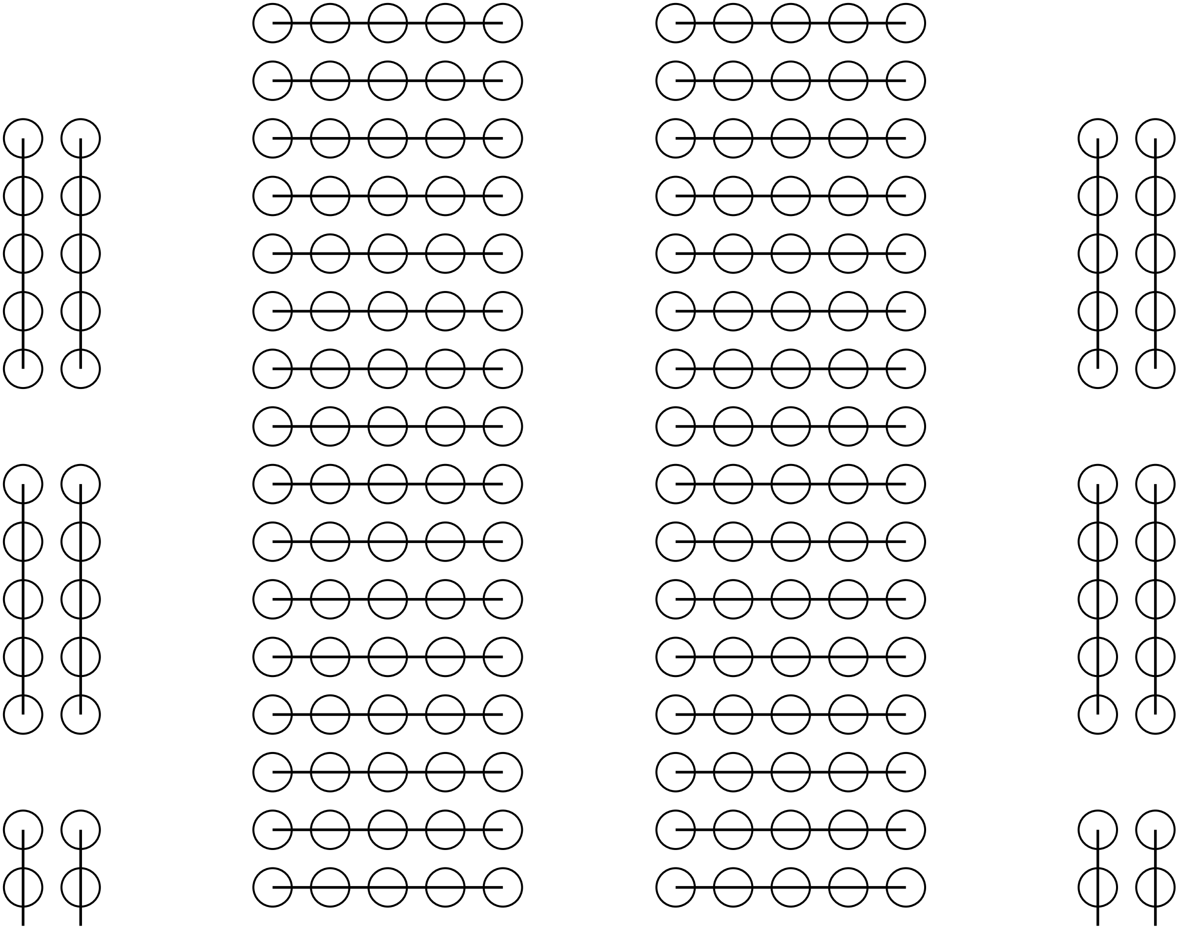
4. Construct the following circuit and use it to determine the internal resistance R_i of each of the 1.5 V batteries.



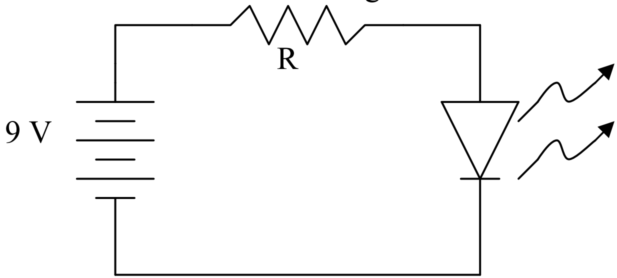
5. Repeat the above determination (5) of R_i for each battery using the motor instead of the light-bulb.

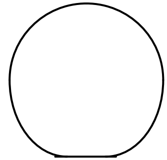
White Boards

Circuits are usually constructed for testing on what is called a “breadboard.” The breadboard is so named because circuits were once soldered together on an actual bread-cutting board, which was a handy way of providing insulation and stability for the circuit. The white board was developed in the 1970s, and is so popular that “white board” and “bread board” are considered synonymous today. The white board consists of sockets that are connected together and into which one plugs various components and wires. The socket connections are illustrated below. The circles represent the sockets, and the lines represent the connections between the sockets. A pair of small needle-nose pliers is helpful for inserting wires and component leads into the sockets. You will also need a pair of wire-strippers for making jumpers from 22 gauge solid wire. Most white boards connect the vertical groups of five in a column together to for what is called a power bus. Some white boards have only some of the vertical groups connected together and others do not connect the vertical groups at all. However the vertical groups are connected, the five sockets forming a group (whether vertical or horizontal) are always connected. Horizontal groups are never connected to other horizontal groups by the manufacturer of the white board. Note the trough, or channel that runs down the middle of the white board: this trough is a very important landmark used when inserting Integrated Circuit devices (ICs) into the white board. (More about this later.)

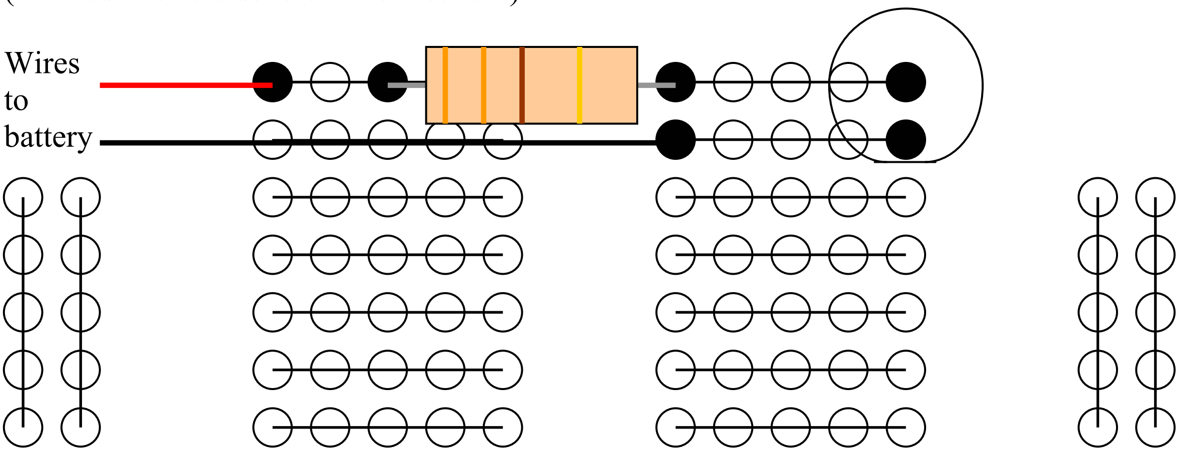


6. Construct the following circuit. R= 330 Ω

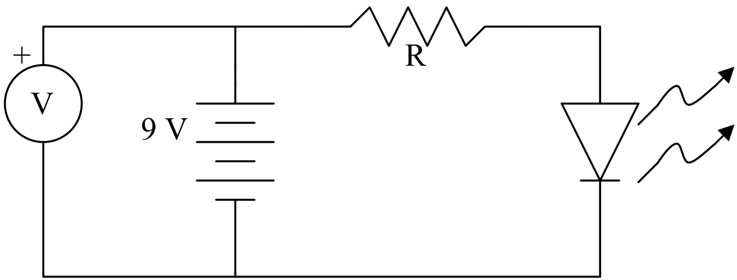




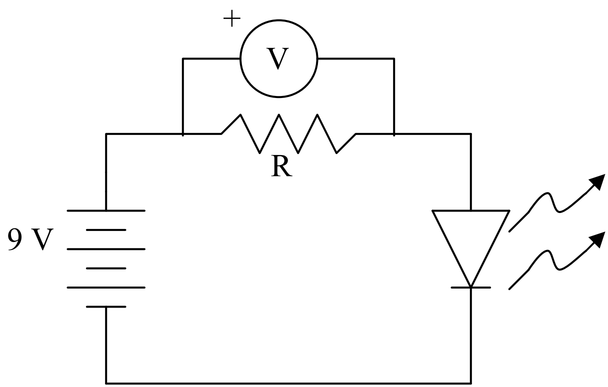
LED packages when viewed from the top have a flat on one side.
The side with the flat is the cathode lead. (Sometimes this lead is shorter.)
Your circuit should look something like the following (top view):
(A filled in circle denotes a connection.)



7. Measure the voltage across the 9 V battery as shown below.



8. Measure the voltage across the resistor as shown below.



From these two measurements, use KVL to calculate the voltage across the LED.
From the measurement of the voltage across the resistor, use Ohm's law to calculate the current through the resistor (and therefore through the LED, since the resistor and the LED are in series). Calculate the power consumed by the LED.

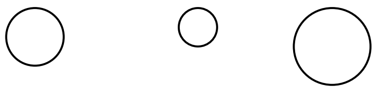
9. Reverse the diode (put it in backwards) and repeat 6., 7., and 8.
10. Repeat 6., 7., 8., and 9. for $R = 1000\ \Omega$.

In your lab report, you should describe each circuit you built with a schematic diagram (not a picture of the white board connections), describe what measurements you made and what the results of the measurement were. Be sure to carry out any calculations required by the exercises.

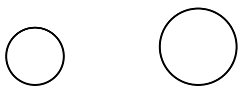
MS Word and Mathcad

To make simple drawings go to View Toolbars and place a check mark next to Drawing by highlighting Drawing and left-clicking. It is important to move the cursor to a point BELOW your proposed drawing with the enter key. Next use the arrow keys to place the cursor at the TOP of where you want to draw a picture. Left-click Draw Grid, change the vertical and horizontal spacing to 0.1 in, and check Display grid lines on screen.

Draw a circle by left-clicking the circle on the Draw tools Tool Bar. The cursor changes to a cross, and you click where you want one edge of the circle to appear, then drag until the circle is of the correct size. (Press esc to remove the message "Create your drawing here.") Note that the circle on the left does not intersect the grid at top and bottom center as do the two circles to the right. This would make it difficult to use the circle on the left as a source in a schematic. Notice that the circles are white not transparent.

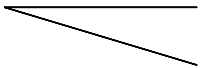


Therefore, if you move the circle on the right over the center circle, the smaller circle is not visible, as shown below.



To remedy this, select the larger circle, choose Format Borders and Shading, move the transparency bar to 100%, and click ok. (Or you may choose Color No Fill.) This is also the place where colors and line widths are chosen.

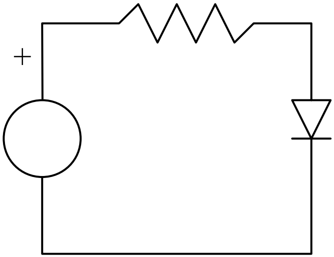
To draw a line click the line on the Drawing tools Tool Bar (the cursor changes to a cross), place the cursor where you want one end of the line, left-click, and drag to where you want the other end of the line.



To draw an arrow click the arrow on the Drawing tools Tool Bar (the cursor changes to a cross), place the cursor where you want to start the arrow, left-click, and drag to where you want the other end of the arrow (the end with the arrow head).

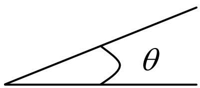
To draw connected lines (like a resistor) use AutoShapes Lines Freeform, then left-click at one and of the first line, release the mouse-button, left-click once where the second line joins to the first, and so on until you are finished at which time double-click the left mouse button.

You can draw simple schematics like the one shown below.



(Points can be moved by choosing an object, and choosing Draw Edit Points)

Curves can be added by choosing AutoShapes Lines Curve, then left-click the first point (release the mouse button), left-click the midpoint (release the mouse button), left-click the last point, then double-click the left mouse button to finish.



Mathematical symbols and Greek letters are most easily entered via the equation editor. To invoke the equation editor use Insert Object Microsoft Equation 3.0 ok
To display the equation

$$\theta = \tan^{-1}\left(\frac{P_W}{Q_{VAR}}\right)^2$$

Select theta from the lower-case Greek menu $(\lambda\,\omega\,\theta)$, press =, then type tan, then choose superscript from the subscript and superscript menu then type -1. Notice the small selection rectangle surrounding the -1: it shows us the location where the next entry will take place. Since we do not want the parentheses to occur in the superscript, press the right arrow key until the entire expression is enclosed, then choose ( ) from the ( ) [ ] and fraction from the fraction roots menu. With the selection rectangle enclosing only the numerator of the fraction type P then choose subscript form the subscript and super script menu and type W. Now use the mouse to move the selection rectangle to the denominator of the fraction, type Q, choose subscript, type VAR, then use the right arrow key to enclose the entire expression in the selection rectangle, choose superscript, type 2, and use File Exit. (Answer yes to save changes.) The equation can be moved with the space bar and the return key in the word editor. The equation size can be changed in the equation editor by selecting Edit Select All, Size Other, and typing the fontsize (in points) that you desire. Before you print, uncheck the Display gridlines in the Draw, Grid menu.

Mathcad

Mathcad is a computer aided design program for doing mathematical calculations. The program is used by assigning numerical values to variable names and then manipulating the variables according to the formulae at hand. For example, suppose you had measured the current through a device as 123 mA and the voltage across the device as 6.4 V, and you wish to calculate the power. You begin by assigning the value 0.123 to the variable I and the value 6.4 to the variable V. The assignment operator in Mathcad is :=
To get this to appear on the screen we simply move the cursor (the red cross) to where we want the definition to appear and type I:0.123 Mathcad adds the equals sign as we type. Next we move the cursor and type V:6.4 then move the cursor and type P:V*I and finally move the cursor again and type P= Mathcad then reports the calculation results. You must remember that the equals sign always tells Mathcad to print the value of a variable or the result of a calculation on the screen. The resulting worksheet is shown below.

$I := 0.123$ $\qquad\qquad V := 6.4$

$P := V \cdot I$

$P = 0.787$

You can put things wherever you want on the screen with the following proviso: assignments of numbers to variables MUST occur above (or in some cases to the left of) the fomula in which they are used. Mathcad can use scientific notation and we could type I:123*10^-3 which would appear on the screen as shown below.

$I := 123 \cdot 10^{-3}$ $\qquad$ $V := 6.4$

$P := V \cdot I$

$P = 0.787$

When typing in Mathcad, keep track of the blue selection rectangle: it tells you where the next entry will go, much like the one in the Microsoft equation editor. (In Mathcad the blue selection rectangle is enlarged by pressing the space bar.) Calculation of a resistance from the values of a voltage and a current would be handled as follows. Let us calculate the resistance of the device whose power we just calculated. We need only add two extra lines as shown below.

$I := 123 \cdot 10^{-3}$ $\quad$ $V := 6.4$

$P := V \cdot I$

$P = 0.787$

$R := \dfrac{V}{I}$

$R = 52.033$

V over I is typed as V/I

Mathcad worksheets are “live” that is if you go back to the top and change any of the variable assignments, the results recalculate immediately, like a spreadsheet.

It is also possible to plot functions in Mathcad. To do so, it is necessary to do two things: define a range variable (This is the variable that will form the horizontal axis.) and define the function. Let us say we would like to plot $120 \cdot \sqrt{2} \cdot \sin(2 \cdot \pi \cdot 60 \cdot t + 45°)$ over two periods, i.e., from t = 0 to t = 34 milli-seconds. To form the range variable, we type t: (which Mathcad renders as :=) then the first value, followed by a comma followed by the second value, followed by a semi-colon (which Mathcad renders as ..) followed by the final value. The second value sets the fineness of the intervals for the plot. We’ll type t:0,0.1*10^-3;34*10^-3 the result is shown below.

$t := 0, 0.1 \cdot 10^{-3} .. 34 \cdot 10^{-3}$

We thus have 341 points for our plot. Before we define the function (which we'll call v(t)), we must change 45 degrees to radians, since Mathcad only understands angles in radians. Since 2π radians = 360 degrees, the required formula is

$$\theta = 45 \cdot \frac{2 \cdot \pi}{360}$$

To get Greek letters in Mathcad, open the view menu and click toolbars then choose Greek. Then type in the definition of θ and of v(t) as shown below.

$t := 0, 0.1 \cdot 10^{-3} .. 34 \cdot 10^{-3}$

$\theta := 45 \cdot \frac{2 \cdot \pi}{360}$

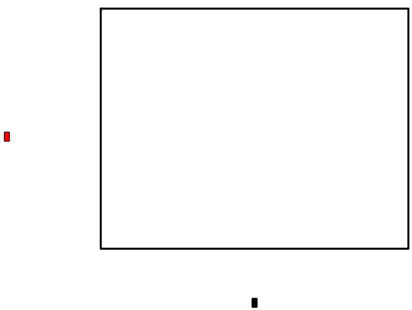
$v(t) := \sqrt{2} \cdot 120 \sin(2 \cdot \pi \cdot 60 \cdot t + \theta)$

Now move the red cursor to a convenient location and open the Insert menu, choose Graph, and choose X-Y Plot. Your work sheet will change as shown on the next page.

$$t := 0, 0.1 \cdot 10^{-3} .. 34 \cdot 10^{-3}$$

$$\theta := 45 \cdot \frac{2 \cdot \pi}{360}$$

$$v(t) := \sqrt{2} \cdot 120 \cdot \sin(2 \cdot \pi \cdot 60 \cdot t + \theta)$$

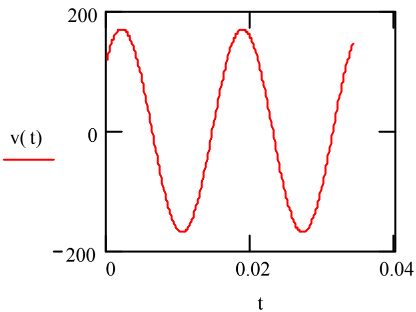


The two squares are place holders. Type t in the bottom one and v(t) in the top one. Your worksheet will change as shown on the next page.

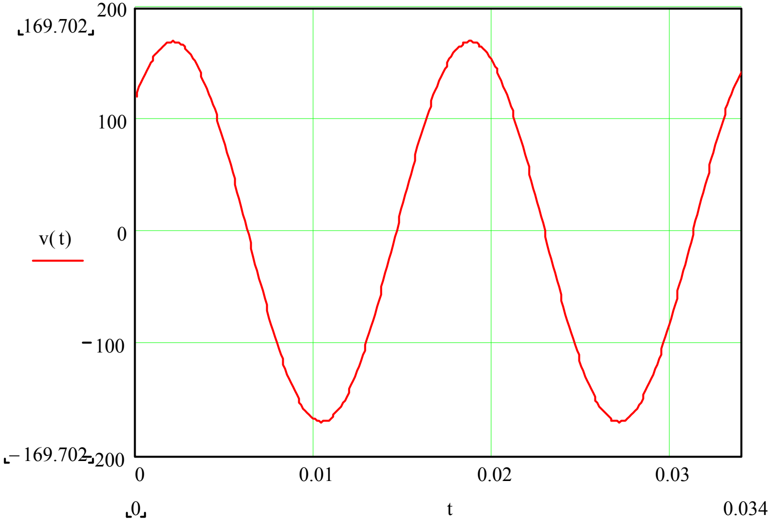
$$t := 0, 0.1\cdot 10^{-3} .. 34\cdot 10^{-3}$$

$$\theta := 45\cdot\frac{2\cdot\pi}{360}$$

$$v(t) := \sqrt{2}\cdot 120\cdot\sin(2\cdot\pi\cdot 60\cdot t + \theta)$$



You can change the size of the graph with the standard windows sizing handle, and you can change many aspects of its appearance by selecting the graph and opening the format menu, choosing Graph, then X-Y Plot.



Notice that $v(t)$ is NOT zero when t is zero: this is the result of the + 45 degrees in the definition of $v(t)$

Tools of the Trade II: Maple

For technical word processing and computer calculation the author now prefers Maple. This powerful Computer Algebra System (CAS), has, in the authors's opinion, better technical word processing and a better drawing editor than MS Word, and a better equation interface than Mathcad. In addition, it's symbolic capabilities mean that you can check any mathmatics you may find in your homework, so that you never need turn in mathematically incorrect homework again!

The first thing we need is to be able to switch back and forth between "math regions" and text regions. This is done with the f5 key. When the cursor is vertical like this | you are in a text region, when the cursor is slanted like this / you are in a math region. Pressing the enter (⟵) key in a text region produces a new line in the usual way. Pressing enter (⟵) in a math region, however enters the current equation into the Maple kernal and gives it an equation number like this (In a math region type ^ to get a superscript, and * to get a multiplication symbol.) :

$x^2 + 2 \cdot x + 1$

$$x^2 + 2\,x + 1 \tag{1}$$

If you find youself at the bottom of the document in a math region (when you press the down arrow the cursor does not move) and you do not wish to enter the equation on the last line into the Maple kernal, simply go to the end of the math region with the right arrow key, press f5, then enter (⟵) several times

to create new blank text lines at the bottom of the document. Then use the up arrow key to put the cursor just below the last equation and press f5 again to change to the math region cursor and continue entering equations.
There are two reasons to enter equations in Maple:

(1) to document a formula, for example:
The power formula for dc excited circuits is $P = V \cdot I$

(2) to manipulate (evaluate, solve, simplify, plot, expand, collect) an equation in Maple. For example:
$P = V \cdot I$

$$P = \mathrm{I}\, V \qquad \textbf{(2)}$$

Note that in the example above the I is not italicized like the *P* and the *V*. This is because Maple uses the capital I as the complex operator, i.e. $I = \sqrt{-1}$. Since electrical and computer engineers often use I and i to represent currents, we tend to use j as the complex operator. This can be achieved by putting the cursor at the beginning of a line in a math region and clicking the [> symbol in the upper tool bar. This produces an isolated math region as shown below.

[>

Put the cursor in the isolated math region, then press f5 and enter the following command:

interface(imaginaryunit=j);

which produces the results below.

```
> interface(imaginaryunit=j);
```

$$I \qquad \textbf{(3)}$$

The blue I indicates the old value of the complex operator. I typically put the above interface command at the top of any Maple document in which I am going to carry out calculations involving currents.

Now when we enter

$P = V \cdot I$

$$P = V\,I \qquad \textbf{(4)}$$

the I is italicized indicating that it is just another unassigned variable. Now we can enter values for any two of the variables and calculate the value of the third. For example,

$eval(\textbf{(4)}, \{V = 12, I = 3.5\})$

$$P = 42.0 \qquad \textbf{(5)}$$

or

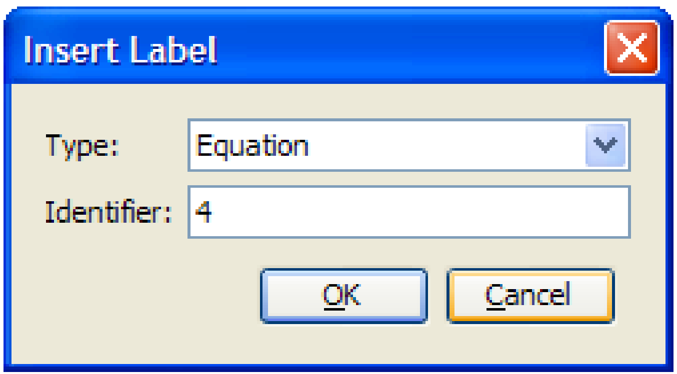
$eval(\mathbf{(4)}, \{P=48, I=3.5\})$

$$48 = 3.5\ V \tag{6}$$

$solve(\mathbf{(6)}, V)$

$$13.71428571 \tag{7}$$

The equation label **(4)** is assigned by Maple when the enter (⟵) key is pressed after entering the equation. To invoke an equation lable, press ctrl- l (i.e., press the ctrl key and the l key simultaneously - this is the "el" key not the "one" key), and a dialog box will pop up and ask you for an equation label.



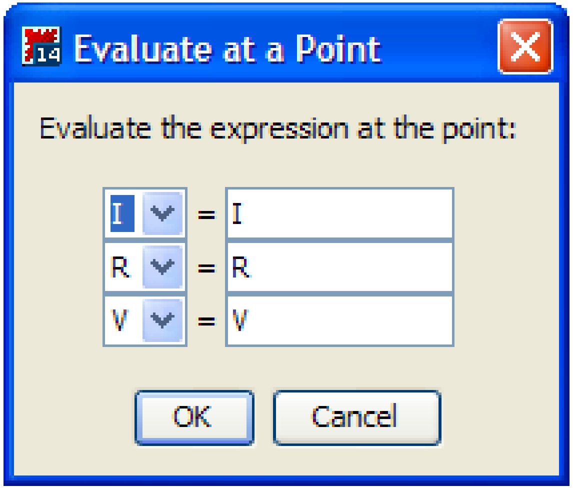
Notice that you cannot simply type (4) to get an equation label - you must use ctrl-l.
Note that in Maple equations are entered first THEN numerical values are substituted into them. This is the opposite of the Mathcad or Matlab interfaces, but much closer to the way one would approach the problem by hand.

It is not necessary to type commands explicitly as above - commands can be selected from the contex sensitive menus invoked by right clicking on the blue version of an equation and selecting "evaluate at a point" from the menu. For instance,

$V = R \cdot I$

$$V = R\, I \tag{8}$$

Right clicking in the blue region of **(8)** and selecting "evaluate at a point" produces the dialog shown below



Filling in V = 12 and R = 6800 gives

(8)

$$V = R\,I \tag{9}$$

$$\xrightarrow{\text{evaluate at point}}$$

$$12 = 6800\,I \tag{10}$$

Note that we can solve **(10)** for I as

$solve(\mathbf{(10)}, I)$

$$\frac{3}{1700} \qquad \mathbf{(11)}$$

It is also possible to click on the blue region of **(10)** and select Solve, Solve for variable I as shown below.

(10)

$$12 = 6800\, I \qquad \mathbf{(12)}$$

$\xrightarrow{\text{solve for I}}$

$$\left[\left[I = \frac{3}{1700}\right]\right] \qquad \mathbf{(13)}$$

which is correct, but in "exact form." There are two alternatives:
(1) Right click on the blue region of **(10)** and select Solve, Numerically Solve as shown below.
(10)

$$12 = 6800\, I \qquad \mathbf{(14)}$$

$\xrightarrow{\text{solve}}$

$$0.001764705882 \qquad \mathbf{(15)}$$

(2)

$evalf(\mathbf{(13)})$

$$[[I = 0.001764705882]] \qquad \mathbf{(16)}$$

Or right click on the blue region of **(13)** and select Numeric Formatting Scientific as shown below.

(13)

$$[[I = 1.7647 \times 10^{-3}]] \qquad \mathbf{(17)}$$

Throughout the rest of this document, we will use the command line form of the commands to get to know the commands and their effects. Virtually all the commands we are going to use can be invoked by right-clicking in a blue region and using the context-sensitive menus.

(1)

$$x^2 + 2x + 1 \qquad \mathbf{(18)}$$

$solve(\mathbf{(1)}, x)$

$$-1, -1 \qquad \mathbf{(19)}$$

$x^3 - 6x^2 + 11x - 6$

$$x^3 - 6x^2 + 11x - 6 \tag{20}$$

solve(**(20)**, x)

$$1, 2, 3 \tag{21}$$

Expand is used to carry out binomial and other multiplications, e.g

$(x-1)\cdot(x-2)\cdot(x-3)$

$$(x-1)\,(x-2)\,(x-3) \tag{22}$$

expand(**(22)**)

$$x^3 - 6x^2 + 11x - 6 \tag{23}$$

Simplify can reduce many expressions, e.g.

$\dfrac{(x-1)}{(x^2-3x+2)}$

$$\frac{x-1}{x^2-3x+2} \tag{24}$$

simplify(**(24)**)

$$\frac{1}{x-2} \tag{25}$$

Collect is used to collect an expression on a variable, e.g.

$expand\left((x+y)^2 + 3 \cdot x \cdot y + (x+3 \cdot y) \cdot (y+2 \cdot x) \cdot (x+1)^2 \cdot (y+2)^2, x \right)$

$$9 x^2 + 33 x y + 13 y^2 + 44 x^3 y + 64 x^2 y + 70 y^2 x^2 + 52 y^2 x + 7 x^3 y^3 + 32 x^3 y^2 + 26 x^2 y^3 + 31 x y^3 + 2 x^4 y^2 + 8 x^4 y + 3 y^4 x^2 + 6 y^4 x + 8 x^4 + 16 x^3 + 3 y^4 + 12 y^3 \tag{26}$$

$collect(\mathbf{(26)}, x)$

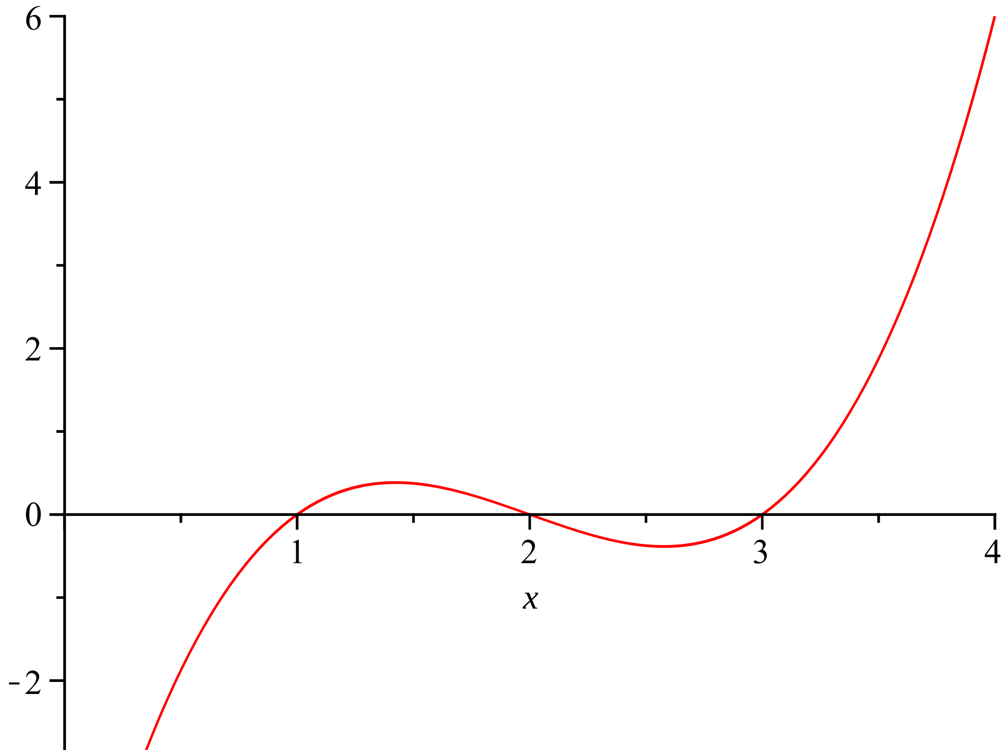
$$\left(2 y^2 + 8 + 8 y\right) x^4 + \left(44 y + 7 y^3 + 32 y^2 + 16\right) x^3 + \left(9 + 26 y^3 + 3 y^4 + 64 y + 70 y^2\right) x^2 + \left(33 y + 52 y^2 + 31 y^3 + 6 y^4\right) x + 3 y^4 + 12 y^3 + 13 y^2 \tag{27}$$

$collect(\mathbf{(26)}, y)$

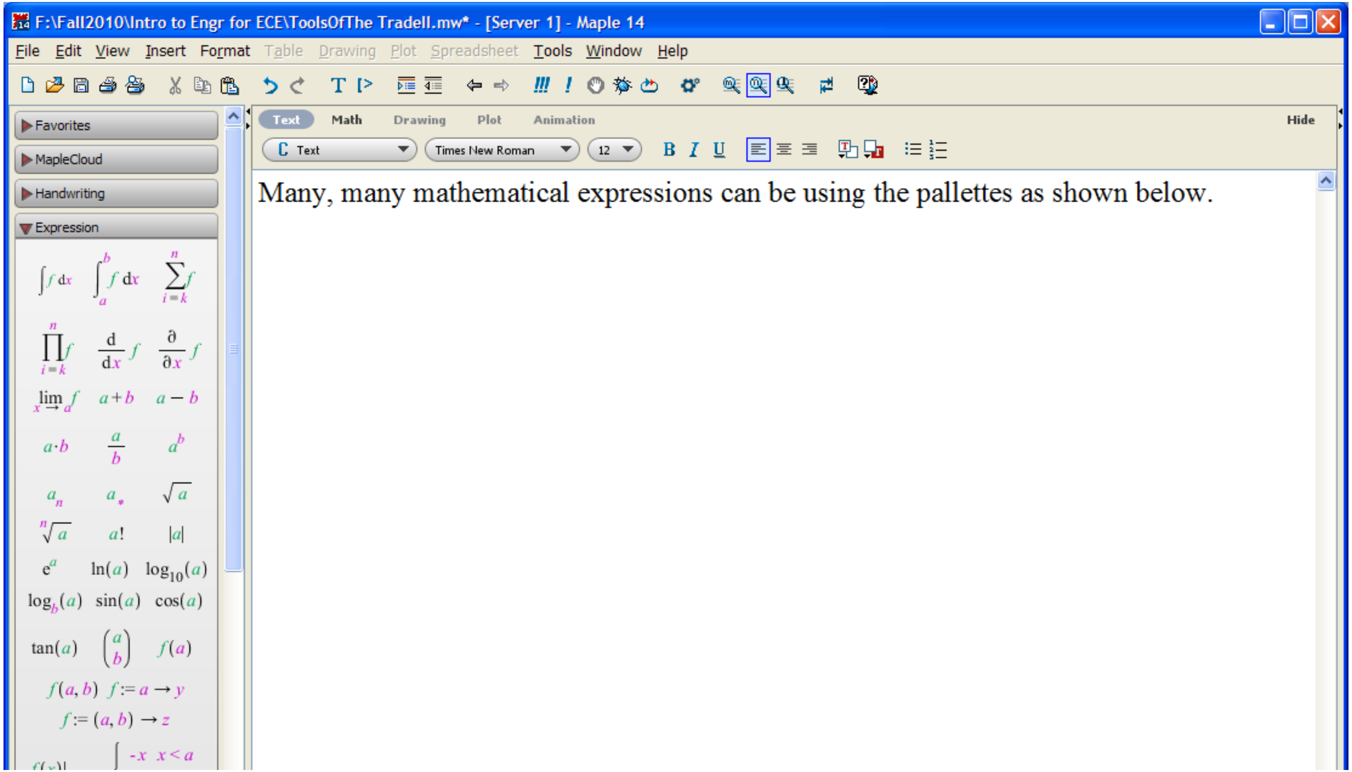
$$\left(3 x^2 + 3 + 6 x\right) y^4 + \left(26 x^2 + 31 x + 12 + 7 x^3\right) y^3 + \left(70 x^2 + 52 x + 2 x^4 + 13 + 32 x^3\right) y^2 + \left(33 x + 64 x^2 + 8 x^4 + 44 x^3\right) y + 9 x^2 + 8 x^4 + 16 x^3 \tag{28}$$

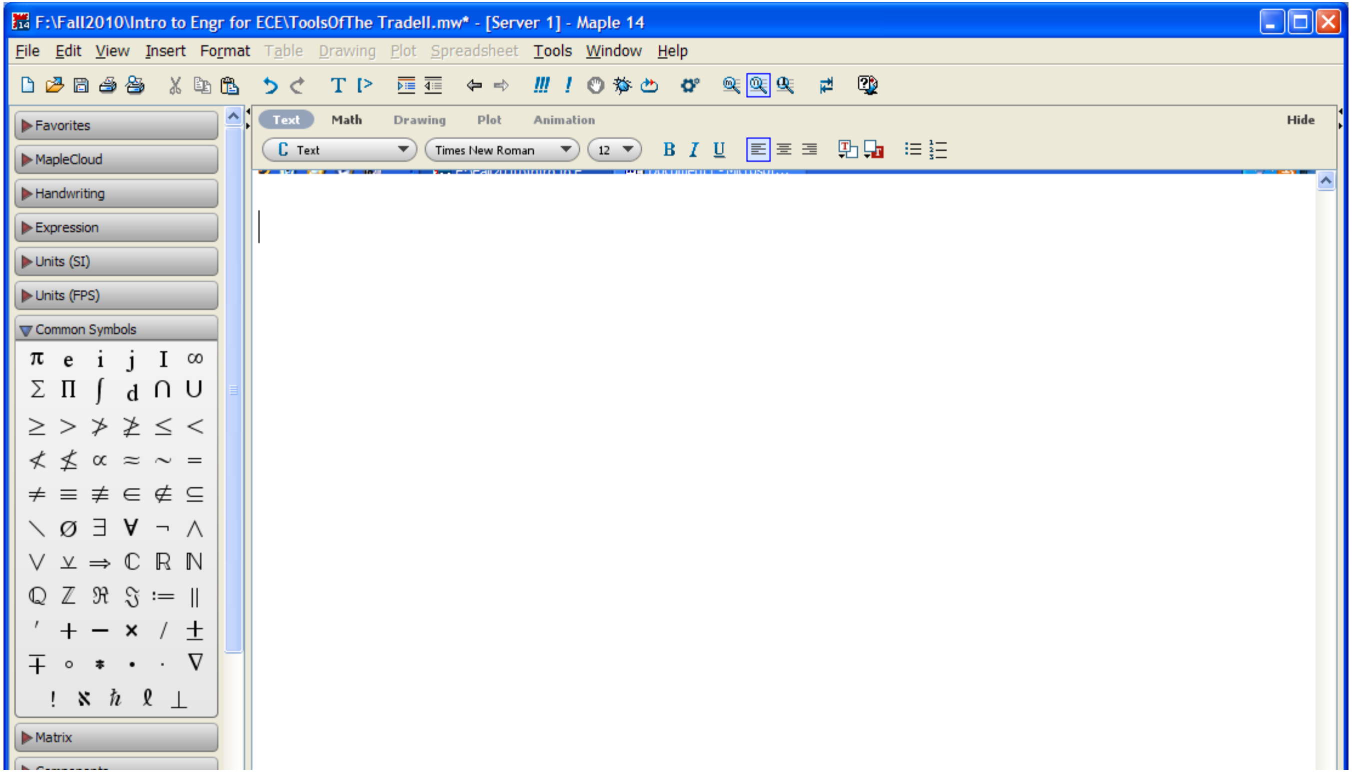
Notice in the plot command below, that the ellipsis used to separate the beginning value of x from the ending value of x is .. (two dots) in Maple rather than the usual three dots. (...)

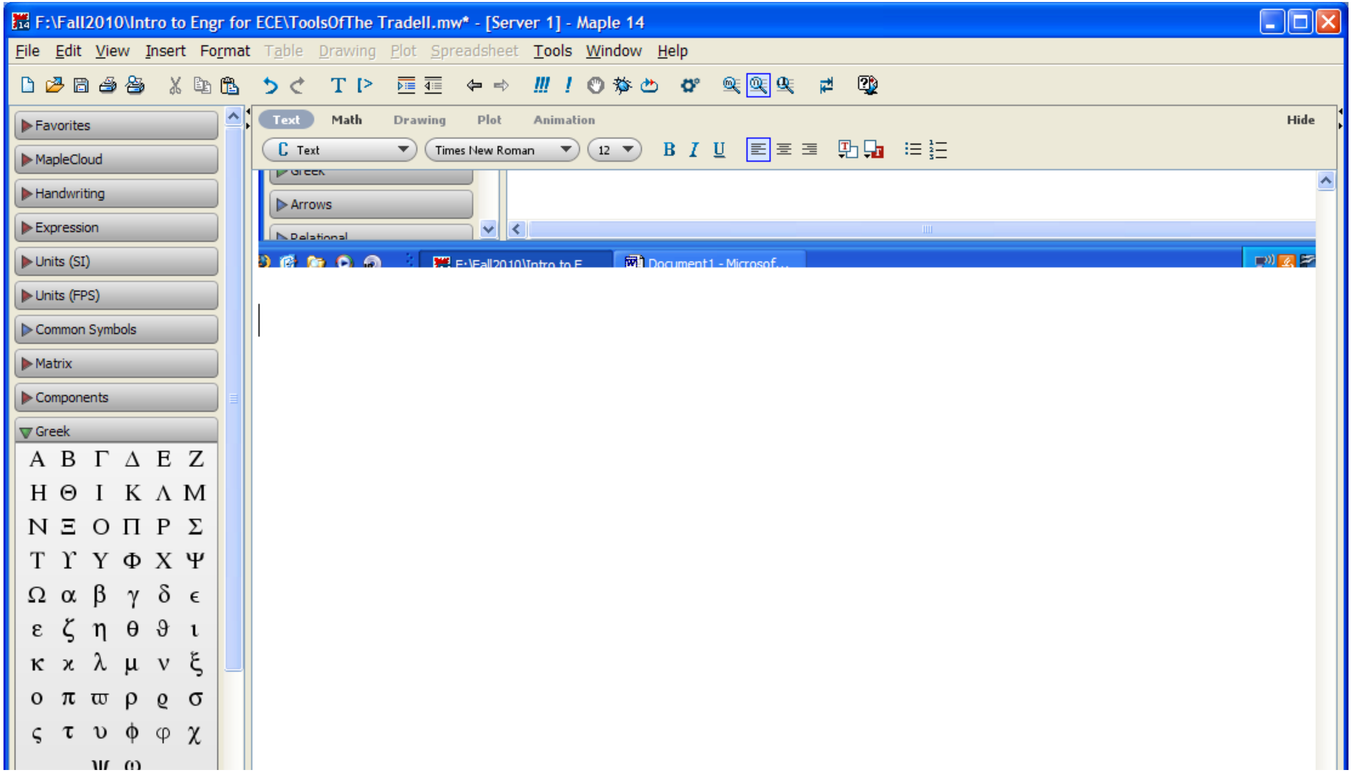
$$plot(\textbf{(23)}, x = 0\,..4)$$

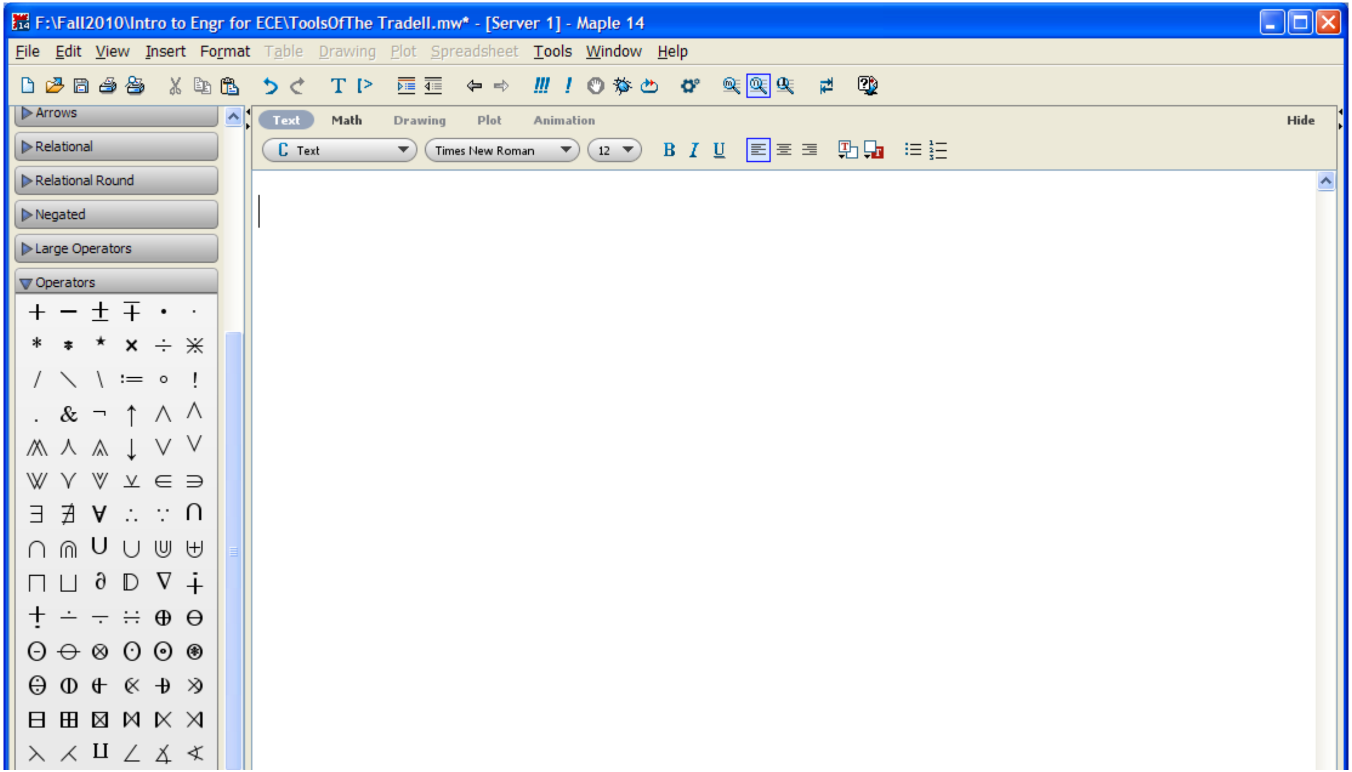

6
4
2
0
-2
1
2
3
4
x

Many, many mathematical expressions can be entered using the pallettes as shown in the screen shots below, where we have opened several of the pallettes.


F:\Fall2010\Intro to Engr for ECE\ToolsOfTheTradeII.mw* - [Server 1] - Maple 14
File Edit View Insert Format Table Drawing Plot Spreadsheet Tools Window Help
Favorites
MapleCloud
Handwriting
Expression
Text Math Drawing Plot Animation
Hide
Text
Times New Roman
12
Many, many mathematical expressions can be using the pallettes as shown below.


F:\Fall2010\Intro to Engr for ECE\ToolsOfThe TradeII.mw* - [Server 1] - Maple 14
File Edit View Insert Format Table Drawing Plot Spreadsheet Tools Window Help
Text Math Drawing Plot Animation Hide
Text
Times New Roman
12
Favorites
MapleCloud
Handwriting
Expression
Units (SI)
Units (FPS)
Common Symbols
Matrix


F:\Fall2010\Intro to Engr for ECE\ToolsOfThe TradeII.mw* - [Server 1] - Maple 14
File Edit View Insert Format Table Drawing Plot Spreadsheet Tools Window Help
Text Math Drawing Plot Animation Hide
Text
Times New Roman
12
Favorites
MapleCloud
Handwriting
Expression
Units (SI)
Units (FPS)
Common Symbols
Matrix
Components
Greek
Arrows
A B Γ Δ E Z
H Θ I K Λ M
N Ξ O Π P Σ
T ϒ Y Φ X Ψ
Ω α β γ δ ϵ
ε ζ η θ ϑ ι
κ ϰ λ μ ν ξ
ο π ϖ ρ ϱ σ
ς τ υ ϕ φ χ
Document1 - Microsof...

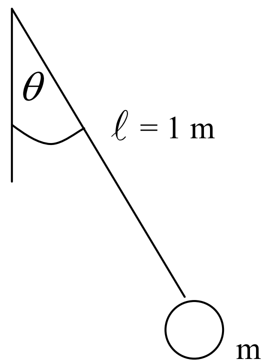

F:\Fall2010\Intro to Engr for ECE\ToolsOfThe TradeII.mw* - [Server 1] - Maple 14
File Edit View Insert Format Table Drawing Plot Spreadsheet Tools Window Help
Arrows
Relational
Relational Round
Negated
Large Operators
Operators
Text Math Drawing Plot Animation
Hide
Text
Times New Roman
12

Sounds, Signals, and Amplifiers

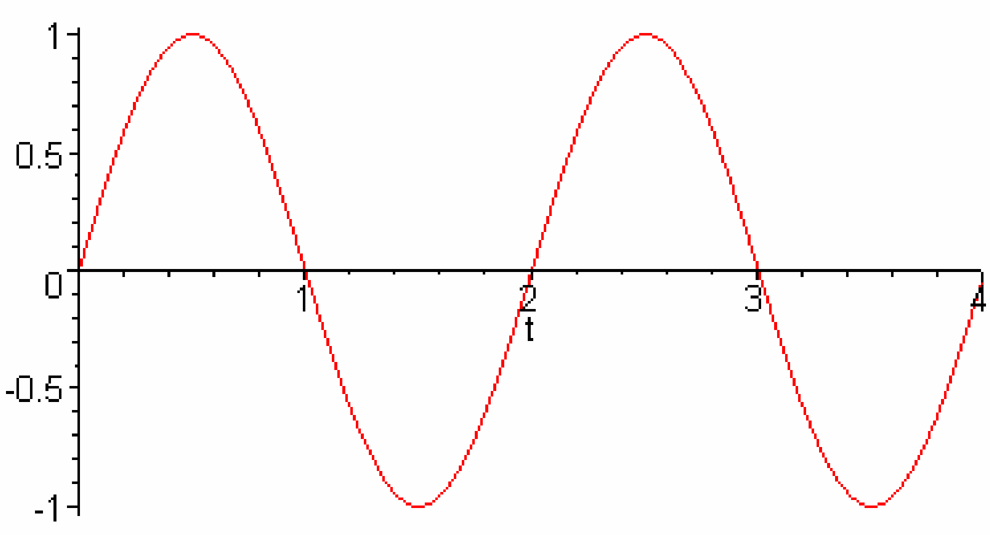
Dr. Richard R. Johnston © 2005

The phenomenon we know as sound is caused by the periodic motion of air molecules. This motion causes changes in air pressure that manifest themselves as pressure waves. Once we have an understanding of the time-varying nature of the air pressure involved we will make a voltage "analog" of the air pressure. This voltage will not be constant as in our battery and LED investigations, but this voltage will be a function of time. Before we investigate sounds let us recall the simplest example of harmonic motion: the pendulum. Harmonic, in this sense of the word, means periodic or repetitive.

If we construct a pendulum as shown below and set it swinging, we find that its speed is not constant but varies with position, and therefore, time.

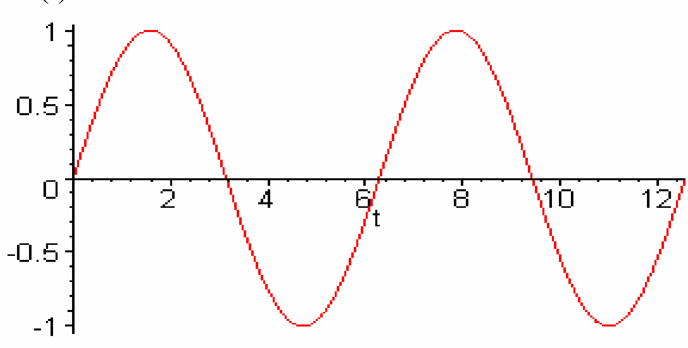


We find that the mass m moves more quickly when the mass is directly under its pivot (i.e. when the angle θ is 0) and more slowly when the mass is further away from the point directly under the pivot (i.e. when the angle θ is larger). The best way to understand time-varying phenomena is to plot them, i.e. to graph the phenomenon on the vertical axis with time as the horizontal axis. If we plot the speed of the mass with respect to time (with t = 0 as the time at which we release the mass) we obtain a plot like the one shown below.
v(t) (The velocity has been normalized to one.)



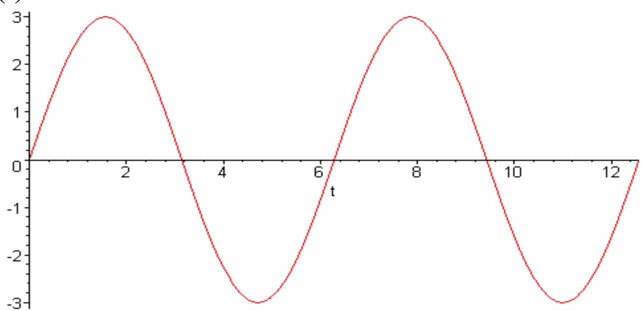
Those who have studied trigonometry recognize this function as $v(t) = k \cdot \sin(\omega \cdot t)$ where ω is a constant, and k is the maximum velocity. We know this isn't sin(t) since the plot of sin(t) is periodic with period 2 pi as shown below. (The plot shown above is periodic with period 2.)

sin(t)



$3 \cdot \sin(t)$ is shown below. (The constant k determines the amplitude of the sine wave.)

3sin(t)



We characterize the pendulum by its period T; for a simple pendulum, T is given by

$$T = 2 \cdot \pi \sqrt{\frac{\ell}{g}}$$

The constant ω in the functional description of the pendulum speed is related to the period T of the pendulum as

$$\omega = \frac{2 \cdot \pi}{T}$$

Note that $2 \cdot \pi \cdot \sqrt{\frac{1}{9.807}} = 2.0064$ so the period of a 1 meter pendulum is about 2 sec. In fact, in the 19th century there was a faction who wanted to define the meter as the length of a pendulum whose half-period is one second, instead of one ten millionth of the distance from the equator to the north pole. Unfortunately small differences in the gravitational force from place to place made a universal definition of the meter from the pendulum period impossible to the desired precision. The variations in the gravitational force are due to the surface irregularities (the earth is not a perfect sphere) and to variations in the density of the earth throughout its interior.

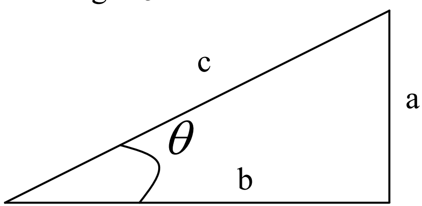
Another way to characterize the periodic motion of the pendulum is to give its frequency in cycles (motion from one point back to the same point) per second. The period t and the frequency f are related as

$$f = \frac{1}{T}$$

The constant ω is called the angular frequency, and the period T, the frequency f, and the angular frequency ω are related as

$$\omega = 2 \cdot \pi \cdot f = \frac{2 \cdot \pi}{T}$$

Any of these three numbers can be used to characterize the period of a harmonic motion. Now those who have studied trigonometry know that the sin function is usually defined in terms of an angle θ as shown below.

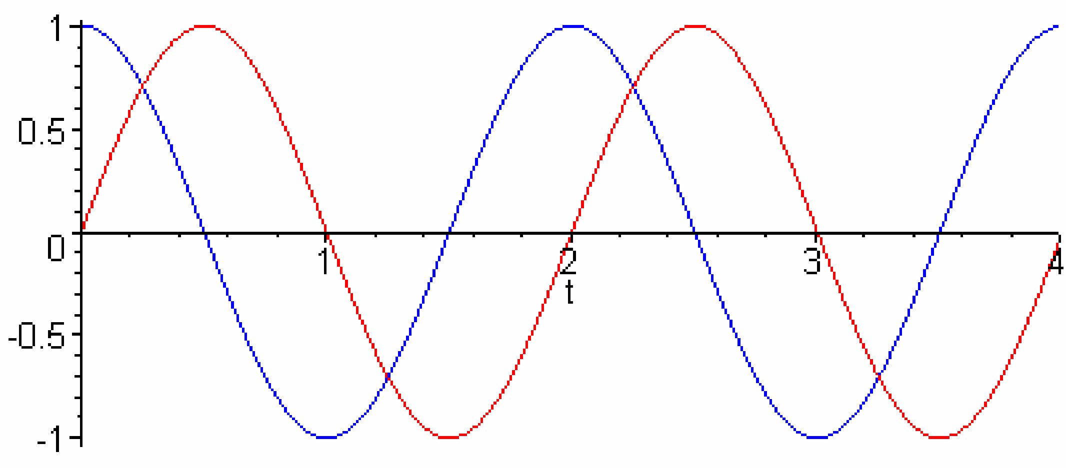


$$\sin(\theta) = \frac{a}{c}$$

We could use this to characterize the POSITION of the pendulum as a function of the angle θ as

$$s(t) = \ell \cdot \sin(\theta)$$

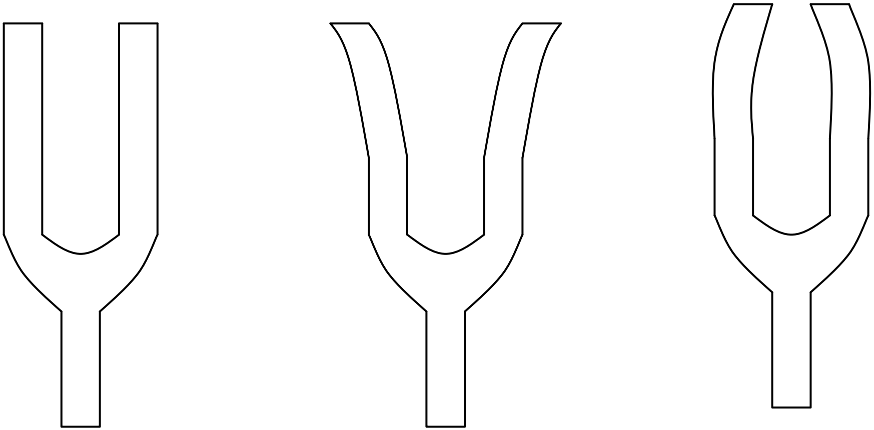
Notice, however, that θ does not equal zero when the time t equals zero. If we plot the velocity and the position both as functions of time we get something like



where the red trace is the speed and the blue trace is the position. (The plots are amplitude normalized to one.) To completely characterize the pendulum (find its maximum velocity and maximum deflection), we would need to know the initial distance that the pendulum was displaced to get it swinging, (We would need the value of the angle θ at time t equals zero.) in addition to the length of the rod. To find the tension in the rod we would also need the mass of the bob.

You will study the pendulum in detail in a physics course; at this point, we are interested in the sinusoidal functions themselves. The functions shown above are said to be 90 degrees out of phase. (Recall that sin $(90^{\circ}) = 1$, and sin $(0^{\circ}) = 0$.)

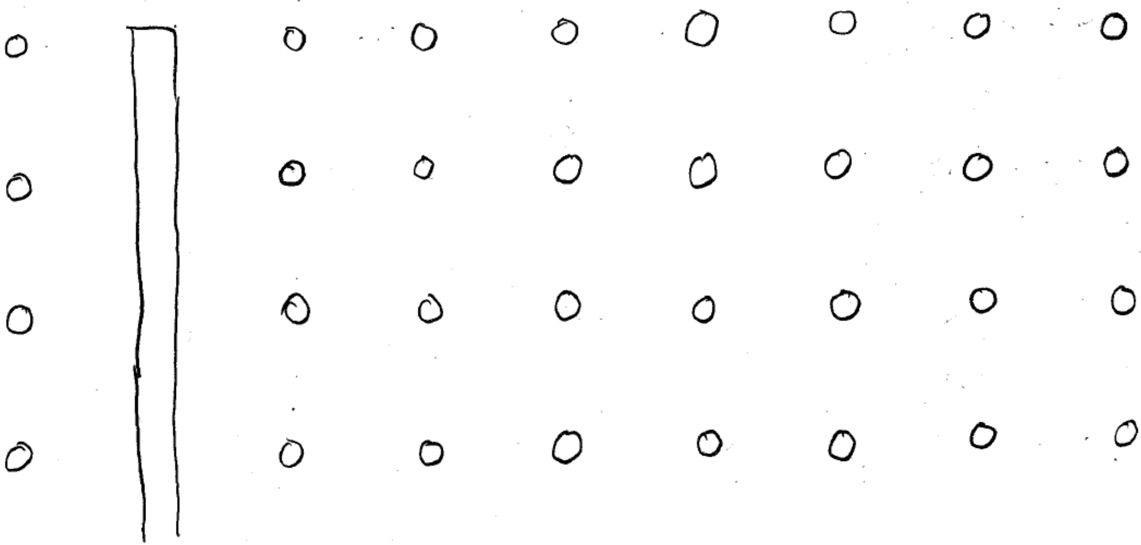
Consider the tuning fork illustrated below.



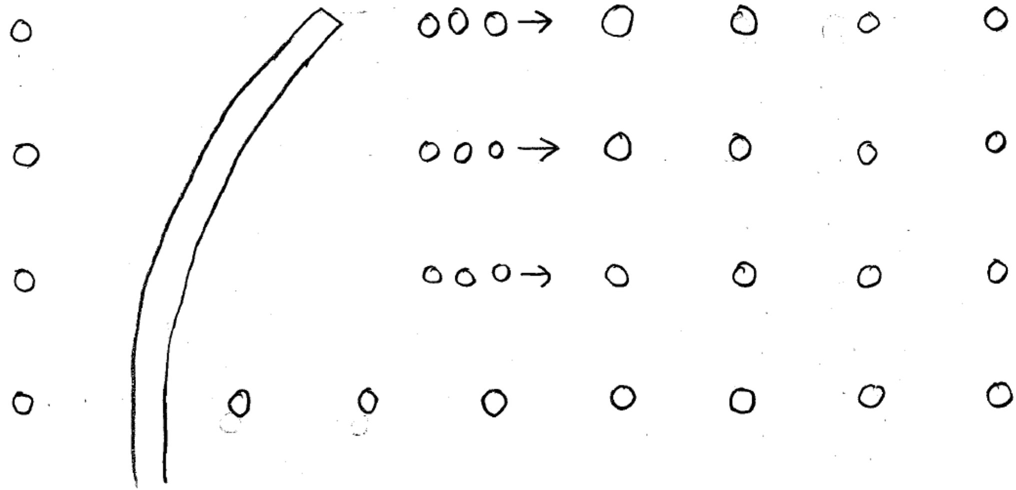
a b c

The fork at a is at rest. When the fork is struck its tines bend back and forth as shown at b and c. When the tines bend out (as in b), they push air molecules together causing a region of higher than normal air pressure; moreover the air molecules are given an outward

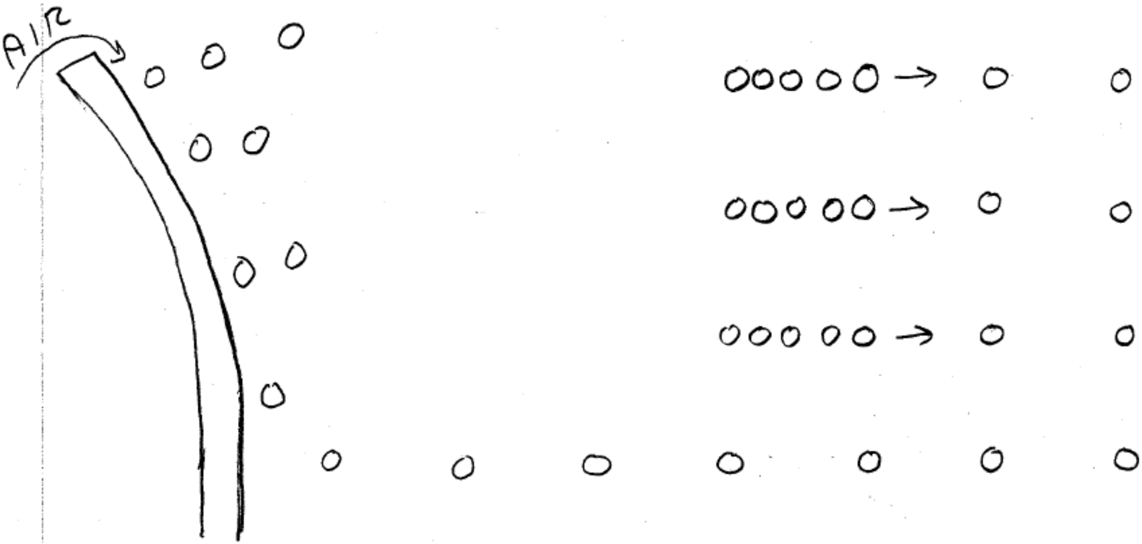
$t = t_0$



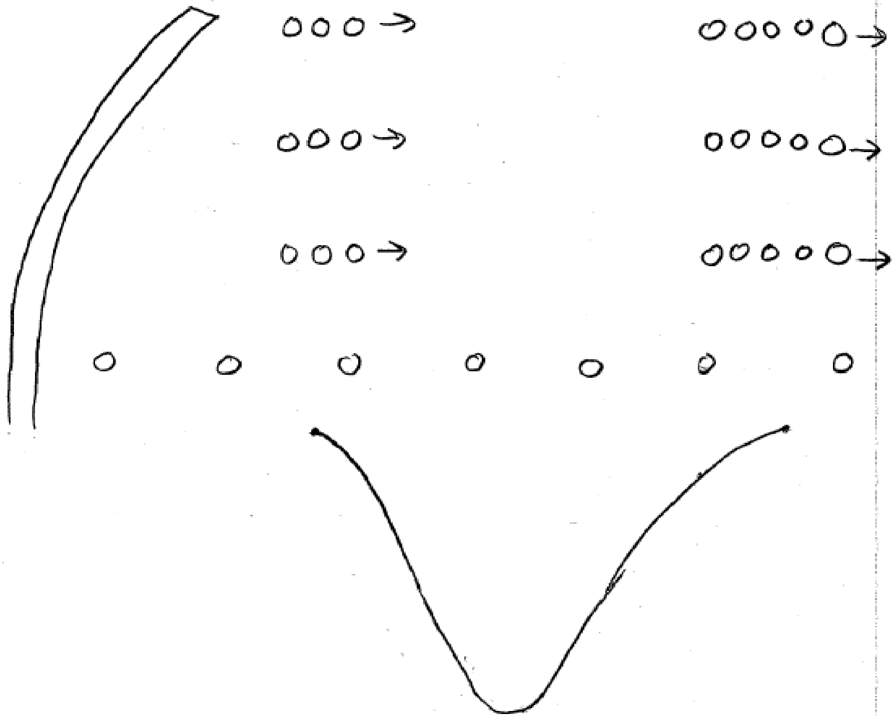
$t = t_1$



$t = t_2$

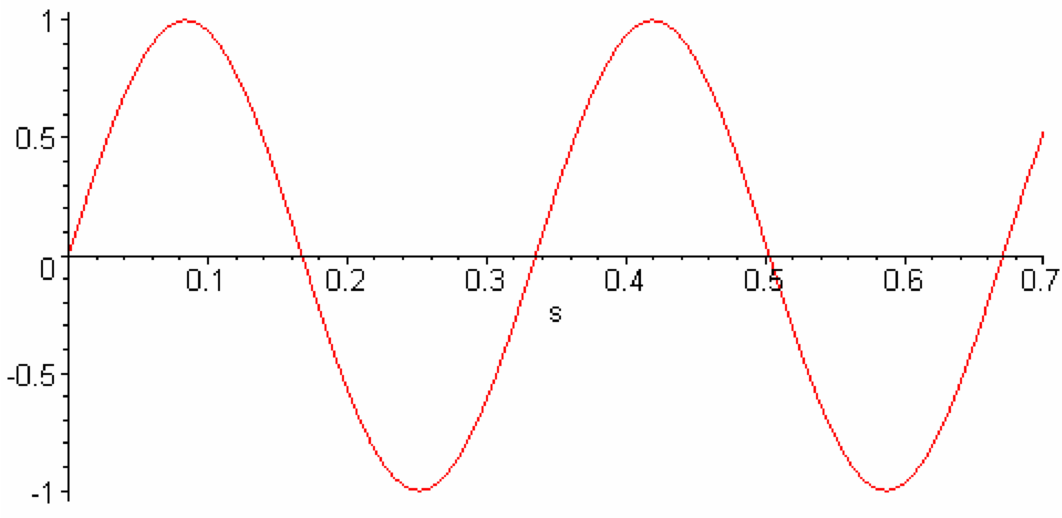


$t = t_3$



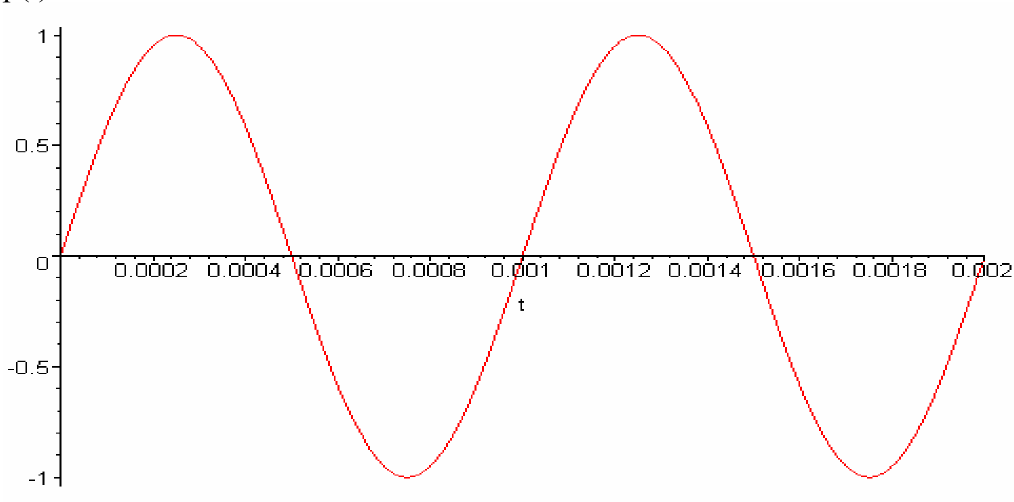
velocity. When the tines bend inward (as in c), they push air molecules inward, leaving a region of lower than normal air pressure. As the tines move back and forth, they cause a train of regions of alternating high and low pressure regions that moves away from the fork. Of course, the air pressure does not change abruptly, but changes gradually. If one plots the air pressure as a function of distance s from the fork, one obtains a plot like the one shown below. (The pressure is normalized to one.)

p(s)



Now since this train of alternating pressure regions is traveling, if we choose a location at a fixed distance from the fork, and measure the air pressure as a function of time we get a plot like the one below. (Notice the relatively short period of oscillation.)

p(t)



This pressure oscillates 1000 times per second. In the late 19th and early 20th centuries this was described as a sound of 1000 cycles per second. During the late 20th century the "cycle per second" was renamed the Hertz (after Heinrich Hertz who did considerable work in electromagnetic waves). Phenomena like sound waves are called traveling waves. We now return to the wavelength of the pressure wave, i.e. the physical length in meters between one high pressure point and the next, or equivalently from one low pressure point to the next, or most commonly from one point of ambient (atmospheric) pressure to the next. Sound travels in dry air at about 335.3 meters per second (1100 feet per second). Let us denote the velocity as v, the frequency as f, and the wave length as λ. Then we have

$$f = \frac{1000 \cdot cycles}{second} \qquad v = \frac{335.3 \cdot meters}{second}$$

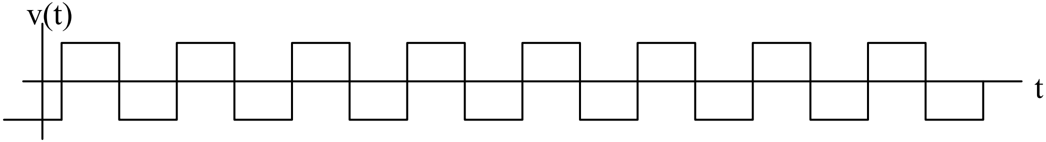
Now we know λ has units of meters per cycle: so if divide v by f, the seconds cancel and we have

$$\lambda = \frac{v}{f} = \frac{\dfrac{335.3 \cdot meters}{second}}{\dfrac{1000 \cdot cycles}{second}} = 0.3353 \frac{meters}{cycle}$$

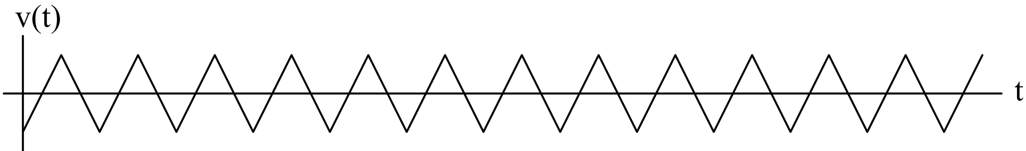
This is an example of a technique called dimensional analysis, (we did the same sort of thing to get units of Watts from V times I in the first chapter) which is very helpful in many physics and engineering problems.
The microphone is a transducer that produces a voltage with the same waveform as the sound wave impinging on it. A transducer is any device that converts the energy in a phenomenon to a time-varying voltage. In addition to sound transducers, there are position transducers, velocity transducers, temperature transducers, fluid pressure transducers, etc.,

etc. The art and science of electronics involves the study and manipulation of time-varying voltages like the one shown above.

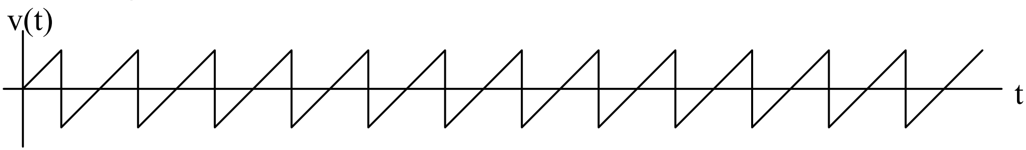
There are two basic pieces of equipment that are indispensable for doing electronics: they are the function generator and the oscilloscope. The function generator (sometimes called a signal generator) produces signals for testing electronic circuits. We will build a rudimentary signal generator later in the course. The oscilloscope is used for examining the wave-shapes of time-varying voltages. This seems an appropriate time to discuss some of the other fundamental electronic waveforms, and their connection to the sine wave.



Square Wave



Triangle Wave



Saw-tooth Wave

There is an amazing theorem called Fourrier's Theorem. (Fourrier should rhyme with "poor today.") The theorem, first suggested by Jean Baptiste Fourrier, says that ANY periodic wave-form (one that repeats over and over) can be represented by a weighted sum of (possibly shifted) sine waves!

$$f(t) = c_0 + \sum_{k=1}^{\infty} c_k \cdot \sin(k \cdot \omega \cdot t + \theta_k)$$

The meaning of this formula is as follows. The $\sum$ is the summation symbol: it means sum the terms to the right, where each variable with a subscript k denote a different term in the summation. The k=1 below the summation symbol, together with the infinity symbol above the summation symbol, mean to let the sum go from k = 1 to k = infinity. So for each value of k (k takes on integer values only) we have a different value of c_k and a different value of θ_k. c_0 is a (possible) dc component. It is possible to think of dc as if it were $\sin(\omega \cdot t + 90°)$ with $\omega = 0$. Mathematically we would say

$$k = \lim_{\omega \to o}\left(k \cdot \sin\left(\omega \cdot t + 90°\right)\right)$$

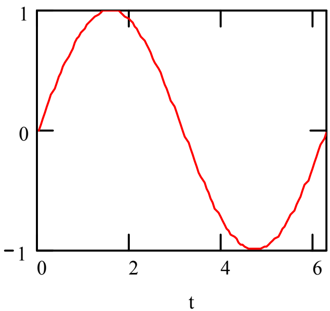
For example, the following graphs show how the addition of sine waves of various frequencies combine to form other wave-forms.

$t := 0, 0.05 .. 2 \cdot \pi$

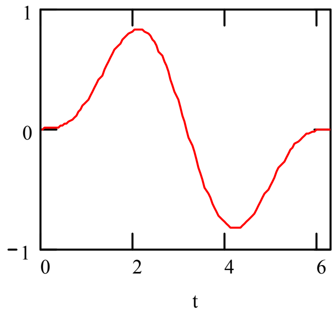
$f_1(t) := \sin(t)$ $\quad f_2(t) := \dfrac{\sin(2 \cdot t)}{2}$ $\quad f_3(t) := \dfrac{\sin(3 \cdot t)}{3}$ $\quad f_4(t) := \dfrac{\sin(4 \cdot t)}{4}$ $\quad f_5(t) := \dfrac{\sin(5 \cdot t)}{5}$

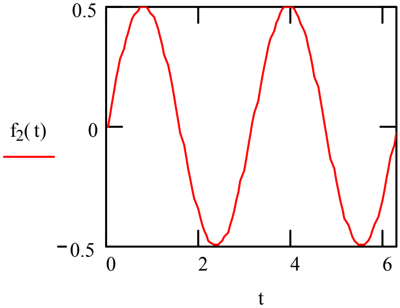
$f_6(t) := \dfrac{\sin(6 \cdot t)}{6}$ $\quad f_7(t) := \dfrac{\sin(7 \cdot t)}{7}$ $\quad f_8(t) := \dfrac{\sin(8 \cdot t)}{8}$ $\quad f_9(t) := \dfrac{\sin(9 \cdot t)}{9}$ $\quad f_{10}(t) := \dfrac{\sin(10 \cdot t)}{10}$

$f_1(t)$

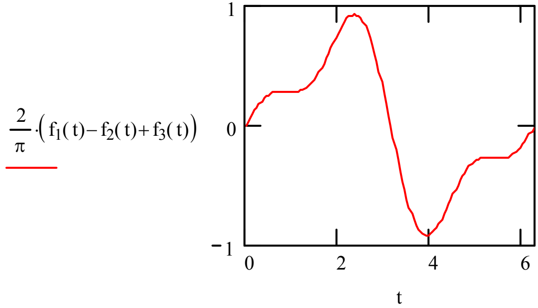


$\dfrac{2}{\pi} \cdot \left(f_1(t) - f_2(t) \right)$

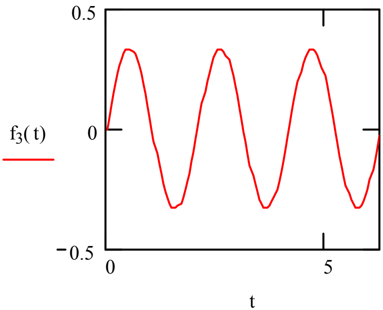




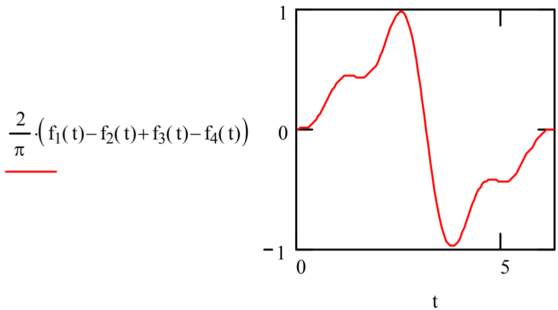
0.5
0
−0.5
0
2
4
6
t
$f_2(t)$



1
0
−1
0
2
4
6
t
$\frac{2}{\pi}\cdot(f_1(t)-f_2(t)+f_3(t))$

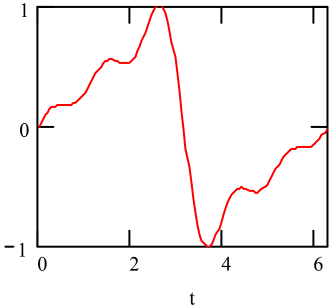


0.5
0
−0.5
0
5
t
$f_3(t)$

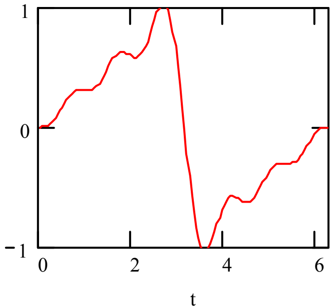


1
0
−1
0
5
t
$\frac{2}{\pi}\cdot(f_1(t)-f_2(t)+f_3(t)-f_4(t))$

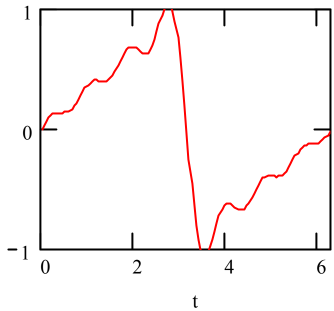
$$\frac{2}{\pi}\cdot\left(f_1(t)-f_2(t)+f_3(t)-f_4(t)+f_5(t)\right)$$



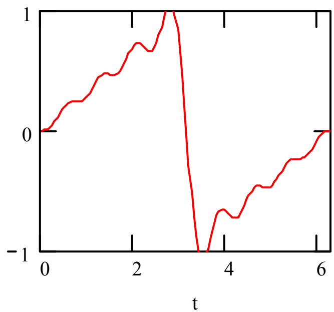
$$\frac{2}{\pi}\cdot\left(f_1(t)-f_2(t)+f_3(t)-f_4(t)+f_5(t)-f_6(t)\right)$$



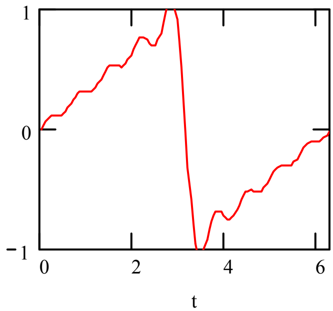
$$\frac{2}{\pi}\cdot\left(f_1(t)-f_2(t)+f_3(t)-f_4(t)+f_5(t)-f_6(t)+f_7(t)\right)$$



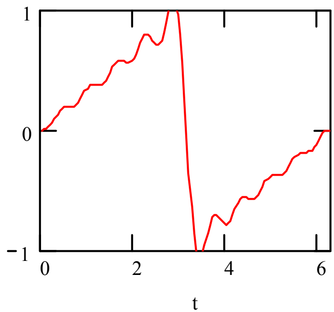
$$\frac{2}{\pi}\cdot\left(f_1(t)-f_2(t)+f_3(t)-f_4(t)+f_5(t)-f_6(t)+f_7(t)-f_8(t)\right)$$



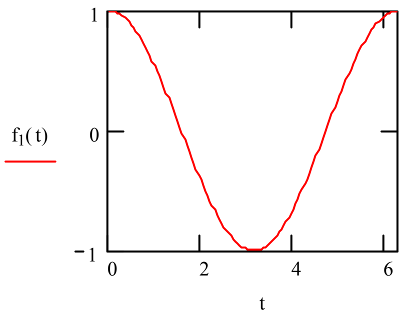
$$\frac{2}{\pi}\cdot\left(f_1(t)-f_2(t)+f_3(t)-f_4(t)+f_5(t)-f_6(t)+f_7(t)-f_8(t)+f_9(t)\right)$$

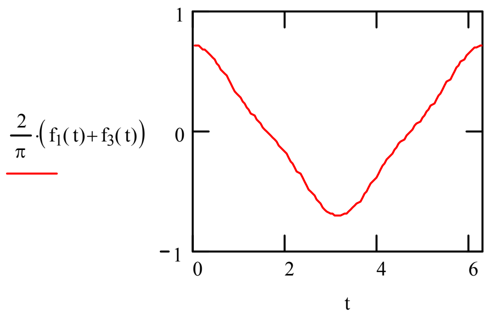


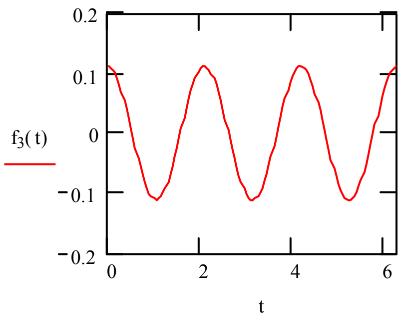
$$\frac{2}{\pi}\cdot\left(f_1(t)-f_2(t)+f_3(t)-f_4(t)+f_5(t)-f_6(t)+f_7(t)-f_8(t)+f_9(t)-f_{10}(t)\right)$$

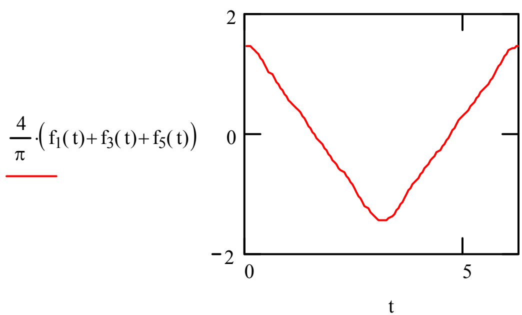


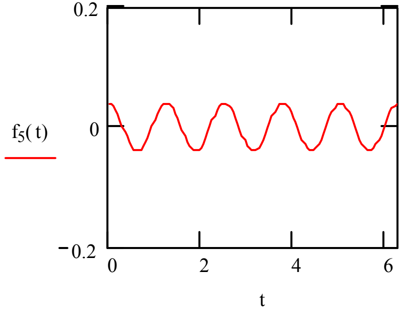
$f_1(t) := \cos(t)$ $\quad f_3(t) := \dfrac{\cos(3\cdot t)}{3\cdot 3}$ $\quad f_5(t) := \dfrac{\cos(5\cdot t)}{5\cdot 5}$ $\quad f_7(t) := \dfrac{\cos(7\cdot t)}{7\cdot 7}$ $\quad f_9(t) := \dfrac{\cos(9\cdot t)}{9\cdot 9}$

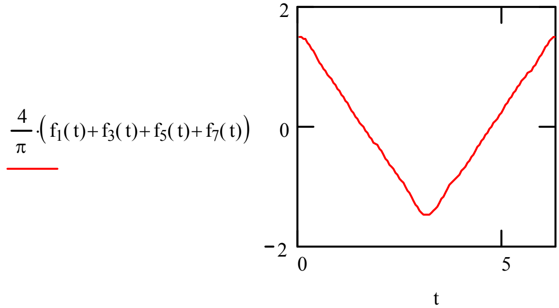


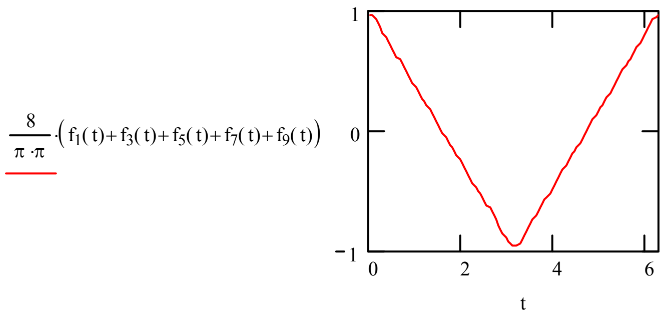




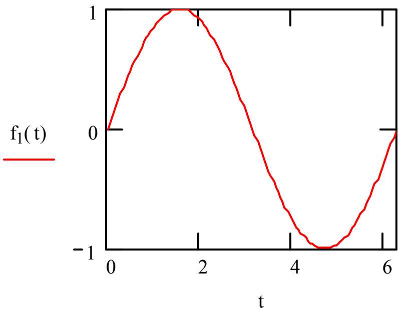


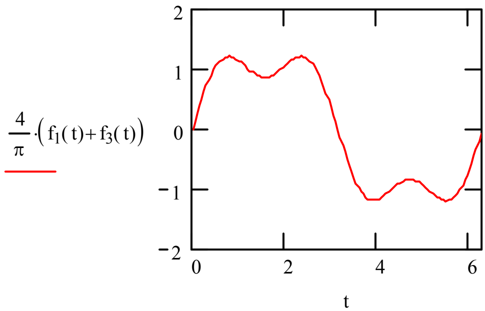

0.2
0
−0.2
0
2
4
6
t
f5(t)

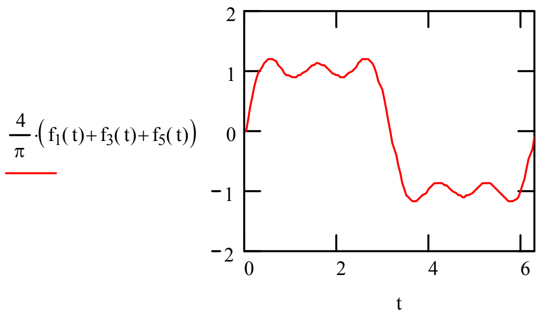

2
0
−2
0
5
t
4/π ·(f1(t)+f3(t)+f5(t)+f7(t))


1
0
−1
0
2
4
6
t
8/(π ·π) ·(f1(t)+f3(t)+f5(t)+f7(t)+f9(t))

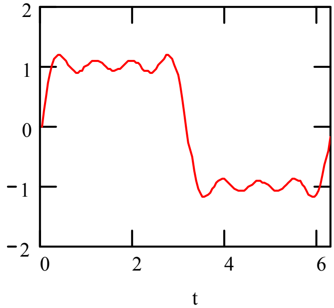
$$f_1(t) := \sin(t) \qquad f_3(t) := \frac{\sin(3\cdot t)}{3} \qquad f_5(t) := \frac{\sin(5\cdot t)}{5} \qquad f_7(t) := \frac{\sin(7\cdot t)}{7} \qquad f_9(t) := \frac{\sin(9\cdot t)}{9}$$



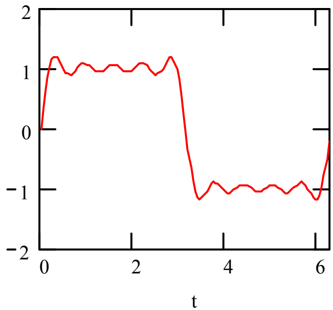




$$\frac{4}{\pi}\cdot\left(f_1(t)+f_3(t)+f_5(t)+f_7(t)\right)$$

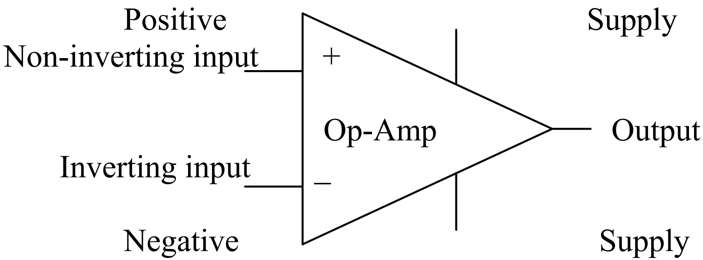


$$\frac{4}{\pi}\cdot\left(f_1(t)+f_3(t)+f_5(t)+f_7(t)+f_9(t)\right)$$



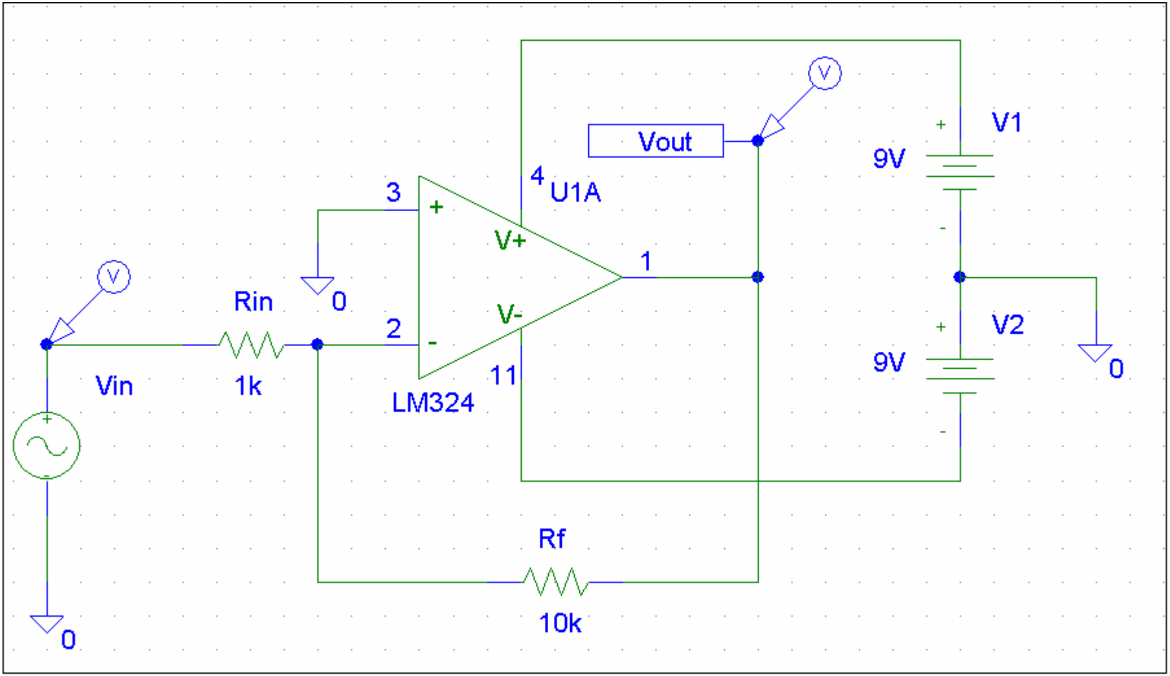
Amplifiers

Most sound sources (electric guitars, CD and tape players, computers, video games, telephones, etc.) produce voltage analogs to sound waves that are too small to drive loudspeakers or head-phones. To change the amplitude of a signal without changing its wave-shape we use an amplifier. An amplifier takes its input voltage wave-form, and produces an exact copy at its output, except that the output amplitude is bigger than the amplitude of the input. The design of transistor amplifiers is fairly complicated and you will study of such amplifiers only after one year of studying electric circuits. There is, however an integrated circuit (IC) called an operational amplifier (op amp) that greatly simplifies the design of many electronic circuits (including amplifiers). One must carefully differentiate between the op amp (the IC) and the amplifier which is built from it.



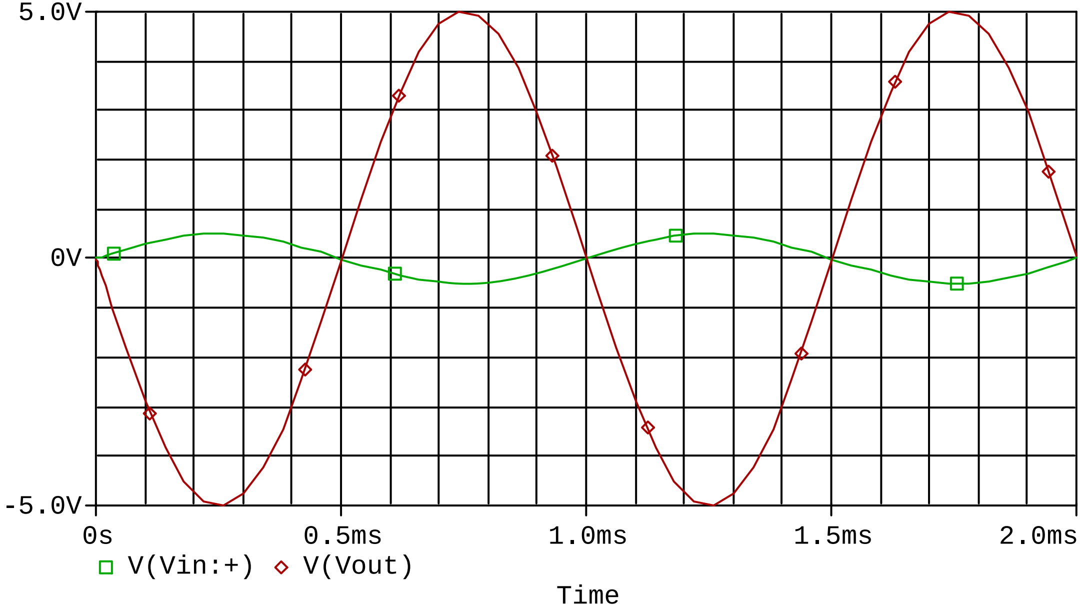
We shall study two different amplifiers based on the op amp: the inverting amplifier, and the non-inverting amplifier.

The inverting amplifier is shown below. The op amp is the IC (U1A), but the amplifier is the circuit consisting of the op amp, the resistors, and the batteries (but not the excitation source). Because Fourrier's theorem assures us that any periodic wave-form can be represented as a sum of sinusoids, amplifiers are universally excited with sine waves for testing and characterization purposes.



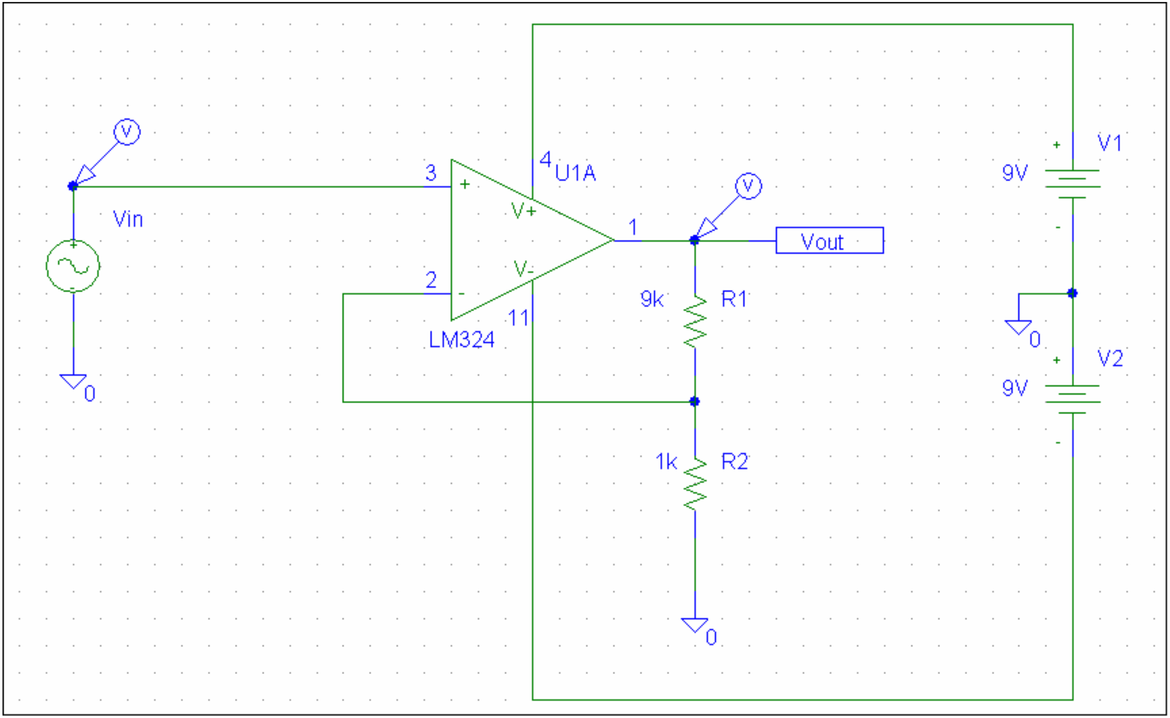
Inverting Amplifier

V_{in} and V_{out} are plotted together below. Note that the output is different from the input in two respects: the amplitude is 10 times larger, and the output is negative when the input is positive, and positive when the input is negative. We call sin waves with this phase relationship as being 180 degrees out of phase. Since the ear is insensitive to steady-state phase differences, this is usually of no consequence.



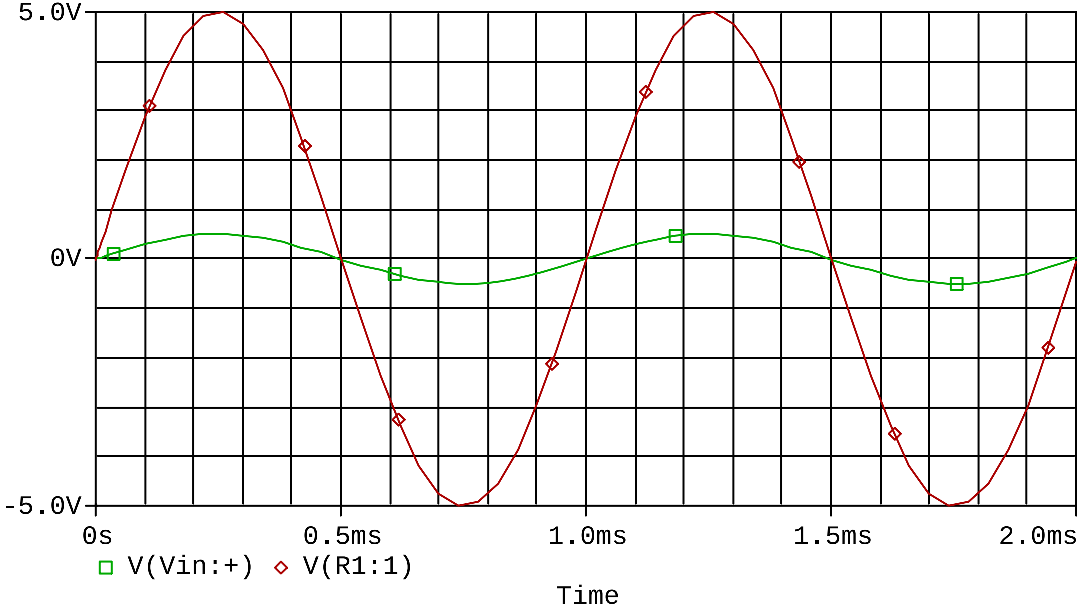
Input and Output Waveforms for Inverting Amplifier

The non-inverting amplifier is shown below.



Non-inverting Amplifier

V_{in} and V_{out} for the non-inverting amplifier are plotted together below.



Input and Output Waveforms for Non-inverting Amplifier

The relationship between V_{in} and V_{out} for the inverting amplifier is

$$V_{out} = -\frac{R_f}{R_{in}} \cdot V_{in}$$

The negative sign signals the fact that V_{out} is 180 degrees out of phase from V_{in}.
We define the voltage gain Av of a circuit as

$$Av = \frac{V_{out}}{V_{in}}$$

so the voltage gain of the inverting amplifier is

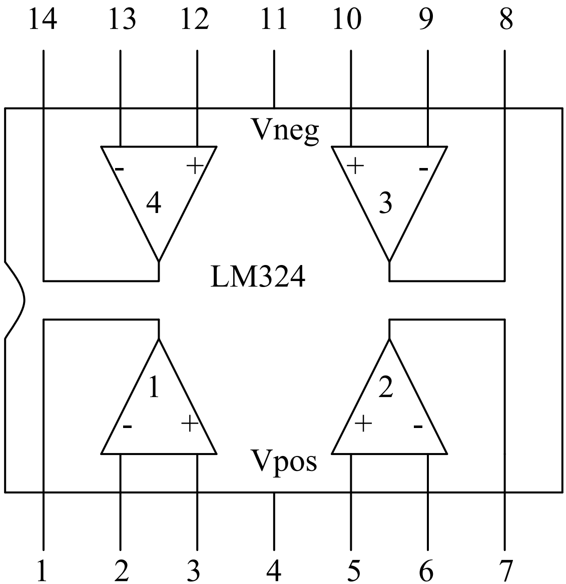
$$Av = -\frac{R_f}{R_{in}}$$

The voltage gain of the non-inverting amplifier is

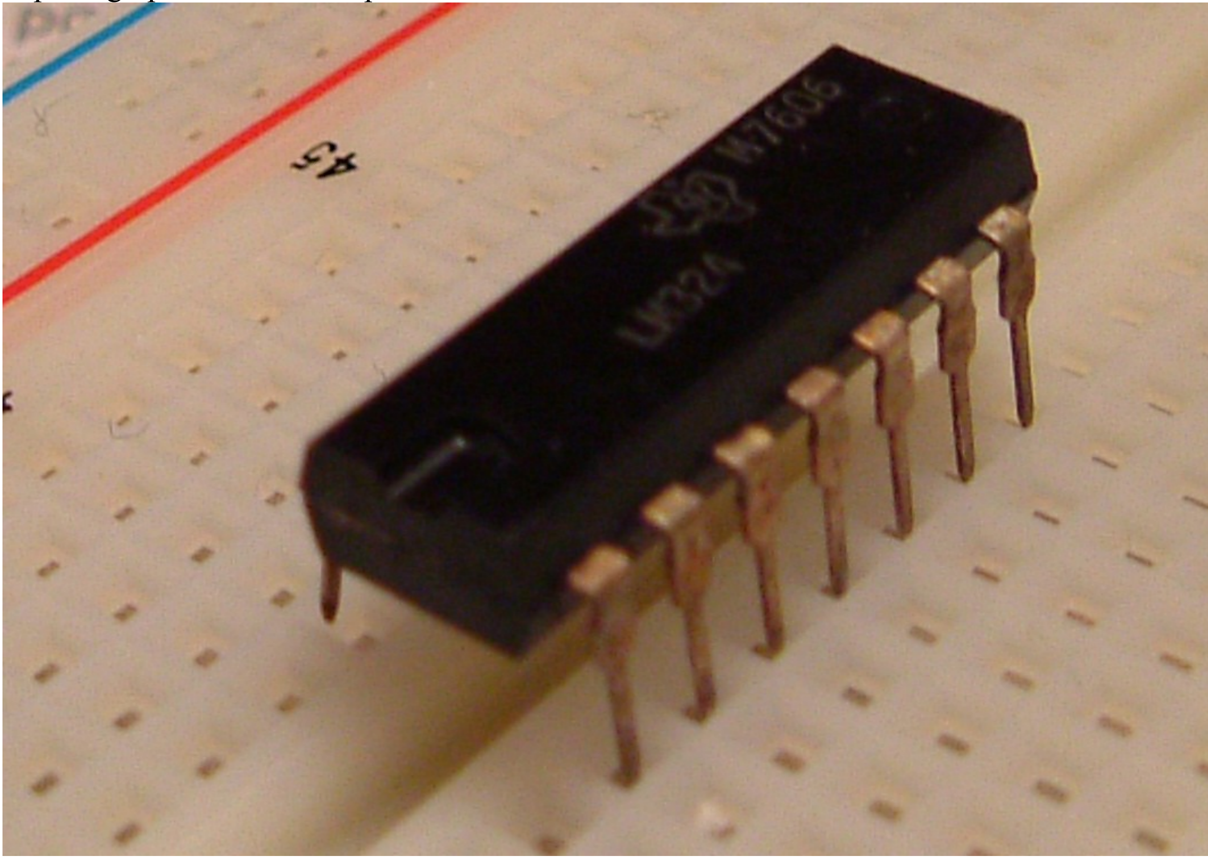
$$Av = \frac{R_1 + R_2}{R_2}$$

Op-amp packages

The op-amp we are using in this course is the National Semiconductor LM324 quad op-amp. This package contains four independent op-amps that share a pair of common power pins. The pins are arranged as shown below. (Diagrams like this are called the “pin-out” of an IC.)



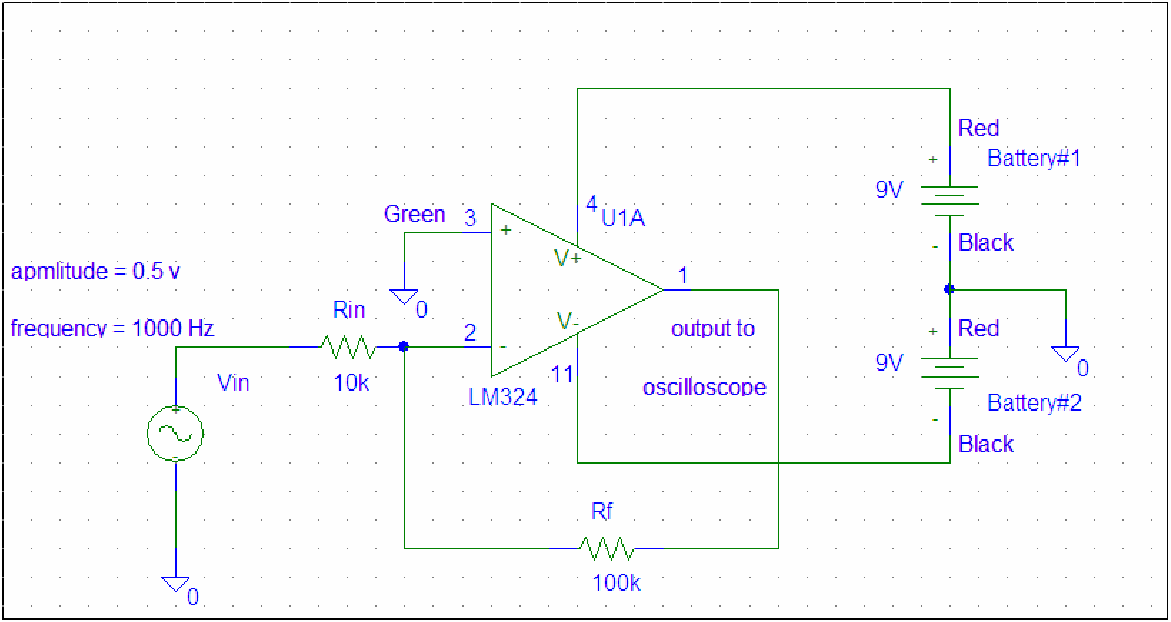
A photograph of an LM324 poised over a white board is shown below.



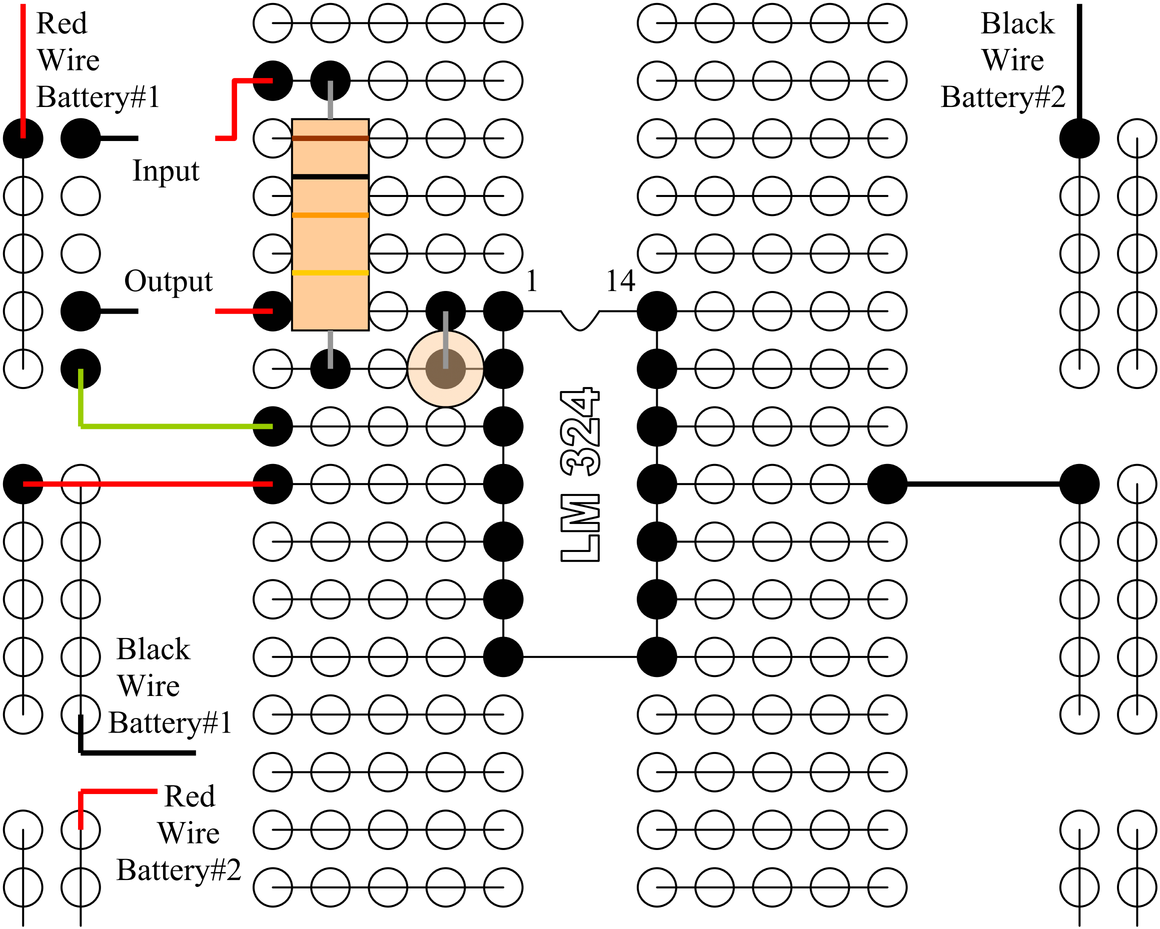
Note that the legs of the IC are slightly wider than the space between the holes on either side of the channel in the middle of the white board. We must, never-the-less, insert the IC into the white board so that it straddles the central channel in the first set of holes on each side of the channel as shown below in the lab exercises. This is done by inserting the leads on one side of the IC SLIGHTLY (not all the way in) into the holes on one side of the channel, then CAREFULLY bending all the leads together until the leads on the other side of the IC are positioned approximately over the holes on the other side of the channel. Now the IC is gently pressed downward into the white board. It is essential to check for bent leads (leads that bent in such a way that they did not enter their respective holes) before proceeding any further.   Failure of one or more IC leads to enter the sockets on the white board is one of the most common reasons for a circuit to refuse to function.

Lab Exercises

#1. Construct the following circuit. (Inverting amplifier)

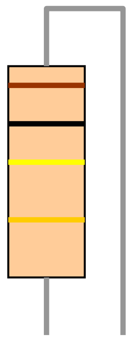


Your circuit should look something like the one shown below. Use the pictorial diagram below as a check – not build directions. The point is to learn how to build circuits from the schematic alone. Future exercises will not show pictorial diagrams!



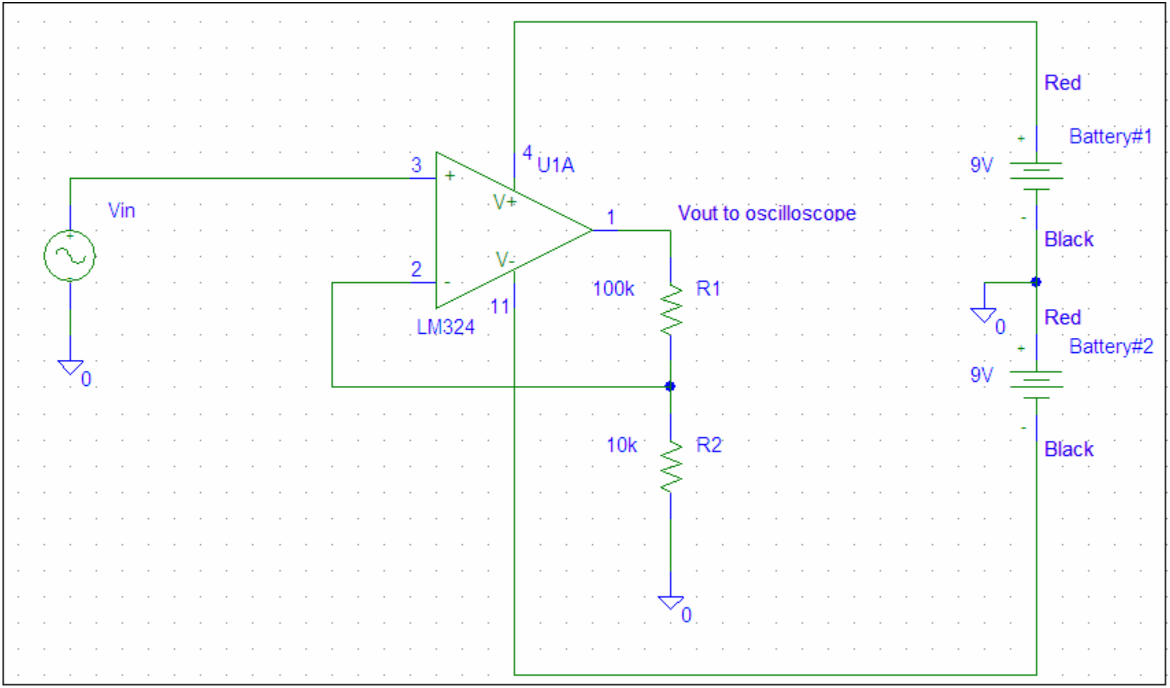
Red
Wire
Battery#1
Input
Output
1
14
LM 324
Black
Wire
Battery#2
Black
Wire
Battery#1
Red
Wire
Battery#2

To get the 100k resistor between pins 1 and 2, bend the leads as shown below, and cut the longer one with your wire striper-cutters.

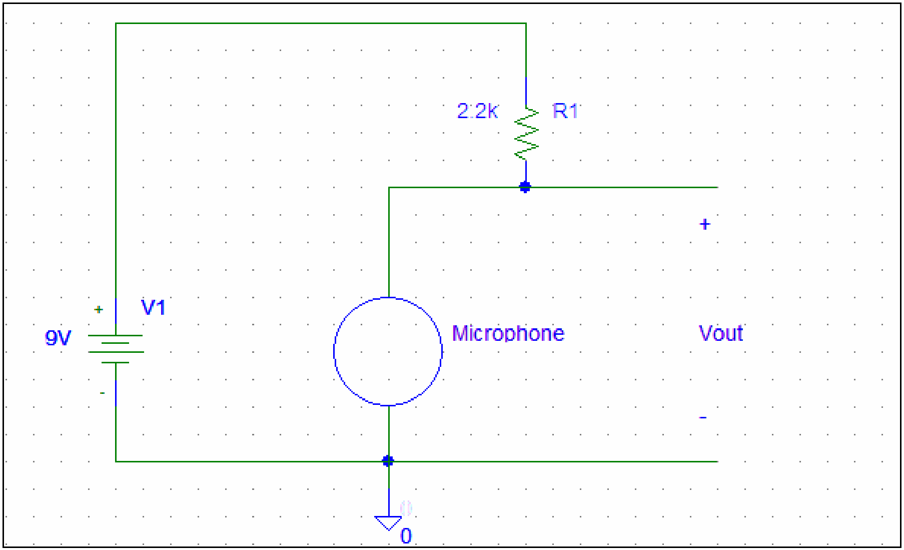


Attach the input to the function generator set at amplitude = 0.5 v, frequency = 1000 Hz, and attach the output to the oscilloscope. View the input on channel one and the output on channel two, triggering from channel one. Print the results for your lab report. In your lab report calculate Av, and describe how your output conforms to your calculation.

#2. Repeat the above measurements for the circuit shown below. (Non-inverting Amplifier)

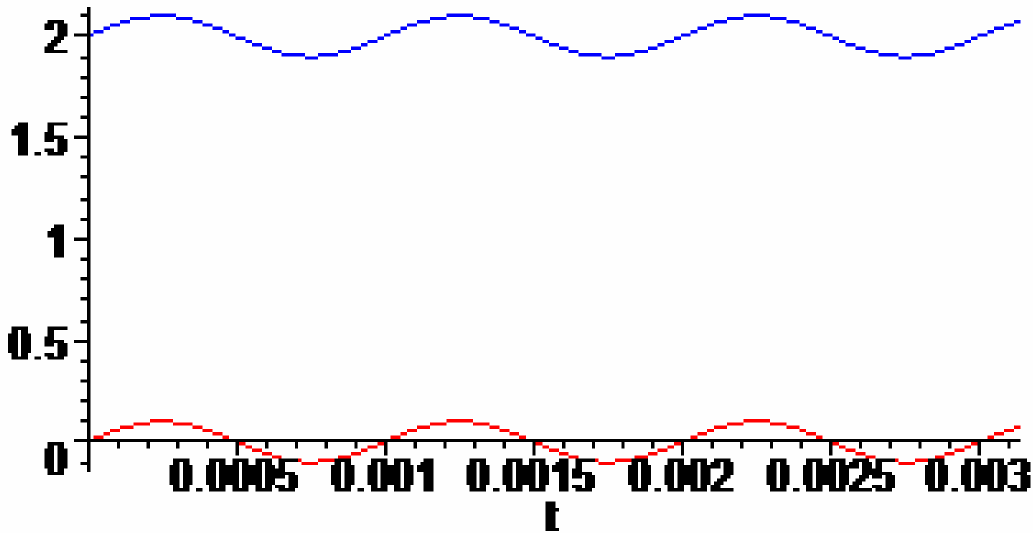


We now wish to connect a microphone and a speaker to the inverting amplifier. Now some microphones require power (usually the application of a dc voltage) and some do not. Carbon button microphones and electret microphones (like the one we are using) require a power supply, while dynamic (permanent magnet) microphones do not. (A dynamic microphone is constructed much like a backwards loudspeaker.) The electret microphones we are using require a 2.2k (in Europe this would be notated 2k2) ohm resistor from the microphone output to the positive dc power supply (9 V in our case) as shown below.

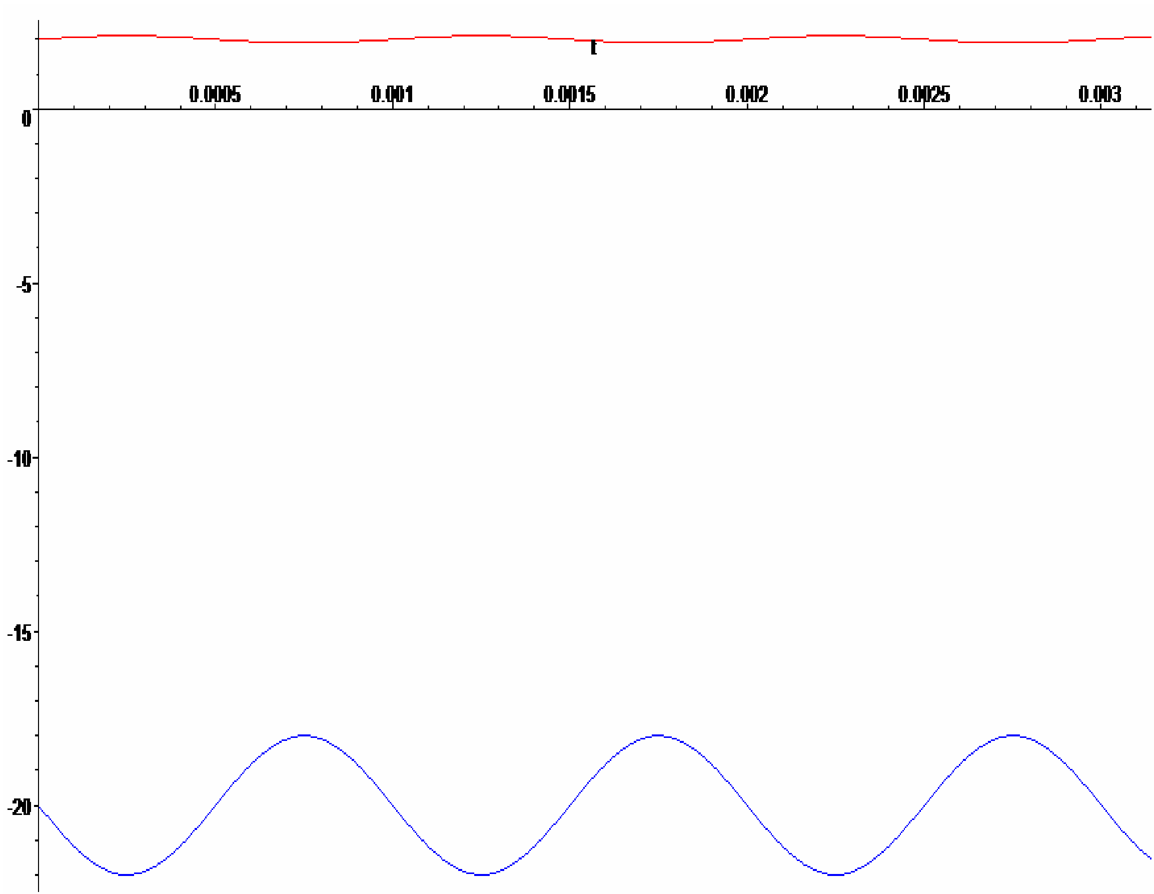


Unfortunately, this arrangement does not give us only the time-varying voltage that is the analog of the sound-pressure waves impinging on the microphone, but a combination of

the time-varying analog voltage we seek, PLUS a dc component. Consider the plots of
$0.1 \cdot \sin(2 \cdot \pi \cdot 1000 \cdot t)$ [the red (or lower) plot]
and $2 + 0.1 \cdot \sin(2 \cdot \pi \cdot 1000 \cdot t)$ [the blue (or upper) plot] shown together below.



Now the problem is that we have a sine wave with amplitude 0.1 that we'd like to amplify so that the amplitude is 1.0. If we apply the blue waveform (the upper wave form) to an inverting amplifier with gain Av = 10 we'd get (theoretically)



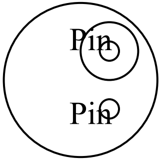
t
0.0005
0.001
0.0015
0.002
0.0025
0.003
0
-5
-10
-15
-20

Where the red (upper) plot is the input to the amplifier and the blue (lower) plot is the output of the amplifier. The difficulty is that our amplifiers have only $\pm 9V$ supplies, which means that the most positive voltage we can obtain with our amplifiers is + 9 V, and the most negative voltage we can obtain with our amplifiers is -9 V. Since the entire output voltage is less than -18 V, the amplifier tries its best and produces a constant -9 V output. (Actually, a little more positive than -9 V, due to limitations of the op-amp.) The amplifier has amplified BOTH the time-varying component of the input waveform (the one in which we're interested) AND the dc component of the input waveform (in which we have no interest). If we wish to amplify signals like the ones produced by our microphones we must find a way to eliminate the dc component of the signal. The ability to pass ac voltages (time-varying voltages) while blocking dc voltages is one of the most important characteristics of a component called a capacitor.

A complete understanding of why the capacitor does what it does in various circumstances, must wait until you have studied sufficient mathematics (Calculus 1, Calculus 2, and Calculus 3) to understand simple differential equations.

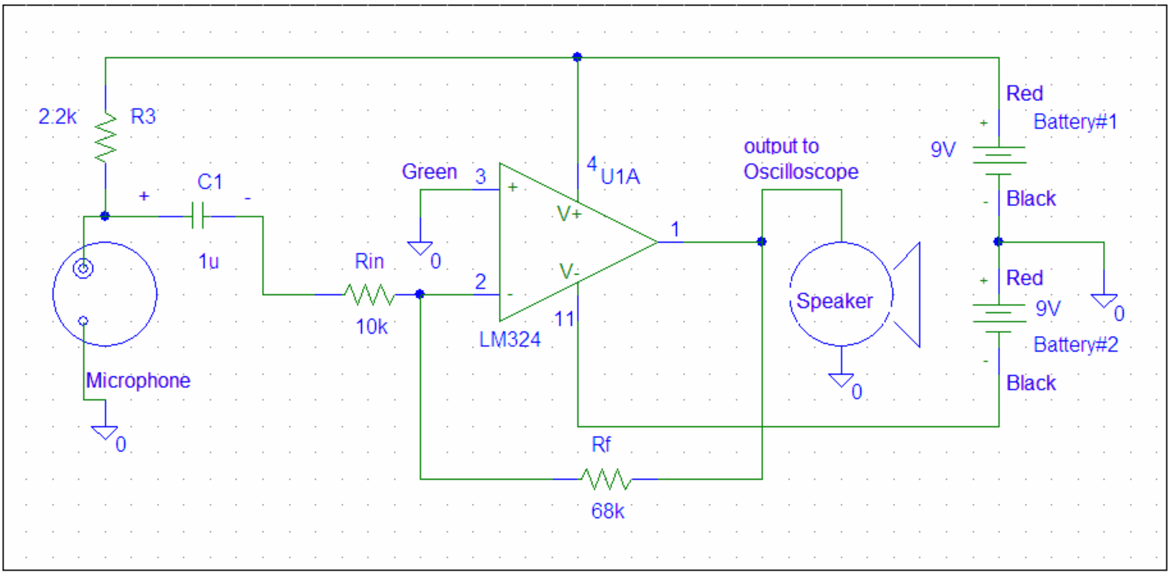
We will build the circuitry necessary to connect a microphone and loudspeaker to our inverting amplifier and analyze the results first, then use a series of experiments to try to understand the behavior of the capacitor in more detail.

#3. Build the circuit shown below and observe and print the voltage waveforms at pin 1 of the microphone, at the junction between the capacitor and the 10k resistor, and at the output (pin 1 of the op-amp) both with and without the loudspeaker in place. Note that one of the microphone pins (pin 1) is isolated from the case (see diagram below) and that the other pin (pin 2) is connected to the case – the polarity is important!

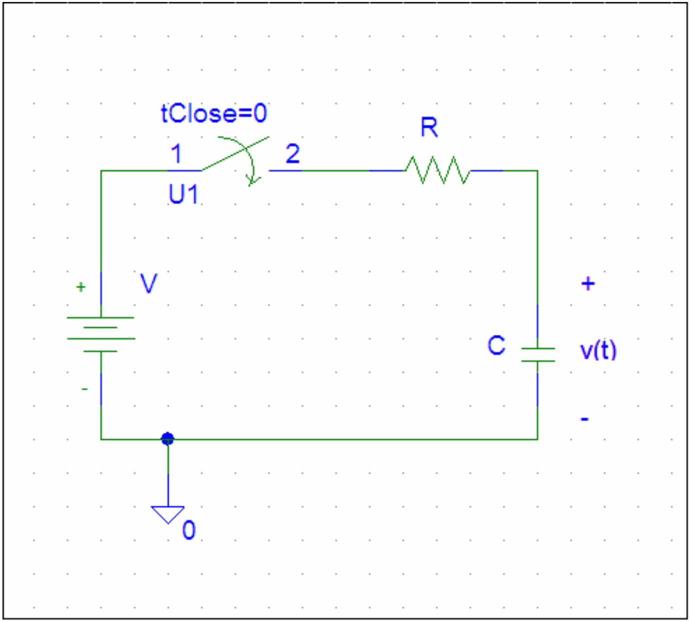


Note that the negative (non-positive) end of the capacitor is marked - - - - - A capacitor inserted backwards can EXPLODE!!!

Microphone (Bottom view)
(Note that the microphone is shown in a top view in the schematic diagram below.)

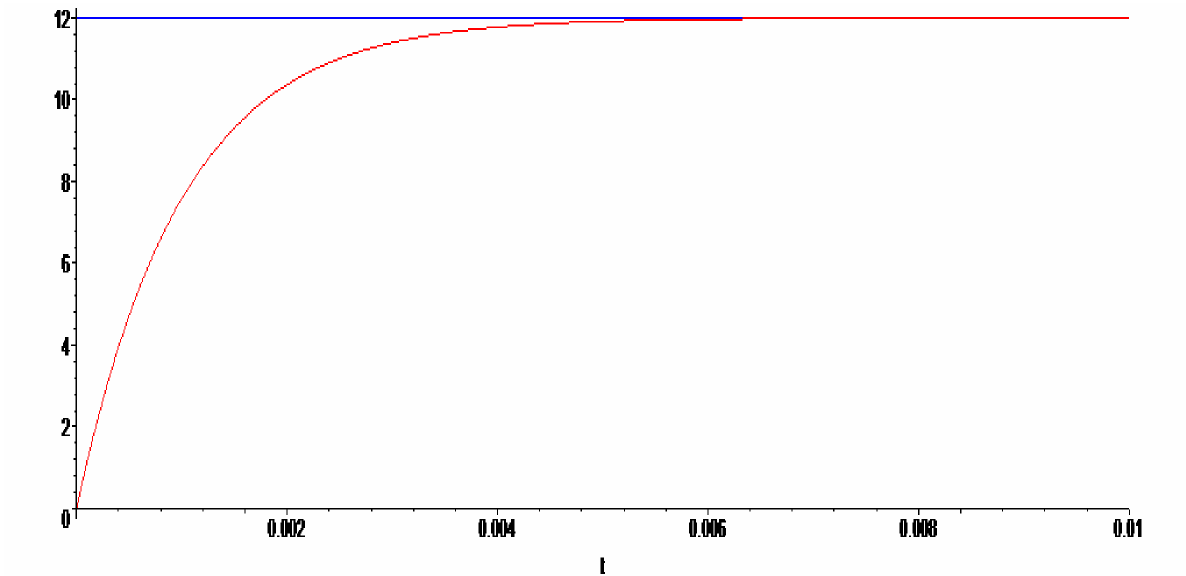


To understand the capacitor's behavior better we shall perform several thought experiments. (Your instructor may demonstrate these experiments to you.)
Let us construct the circuit shown below.



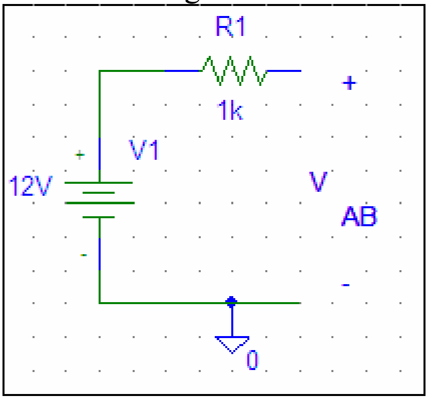
If we close the switch at $t = 0$, we obtain the following response.

$v(t) = V \cdot \left(1 - e^{-\frac{t}{R \cdot C}}\right)$ where e^x is the exponential function and e is the base of the natural logarithms, i.e. $\ln(e) = 1$ and e = 2.718281828. The results for the case V = 12 V, R = 1000 ohms, C = 1 microfarad are plotted below.



Notice that after about 5 msec, the capacitor voltage is about equal to V.

The quantity $R \cdot C$ is called the time-constant τ, and for the case where R = 1000 ohms and C = 1 microfarad, $\tau = 0.001$ The convention is to say that the capacitor voltage stops changing after 5 time constants: we call a capacitor fully charged after 5 time constants. Notice that a <u>fully charged</u> capacitor behaves exactly like an open circuit to the dc source. The current through the resistor is zero and the voltage across the capacitor is the same as the voltage across the source. This surprises most students who expect the voltage between the right side of the resistor and ground (V_{AB}) in the circuit below to be zero.



However, a brief review of Kirchoff's circuit laws should convince us that V_{AB} in the above circuit is 12 V **NOT** 0 V.   KCL requires that the current in the resistor is zero because the current in the open circuit is zero: in fact, an open circuit is defined to be an element whose current is zero for all time, no matter what the voltage across the element may be. But if the resistor current is zero, then the resistor voltage must be zero, and KVL requires that the resistor voltage plus V_{AB} must add up to 12 V. But if the resistor voltage is zero, then V_{AB} = 12 V. This result is central to all of electrical and computer engineering: you must know both the result, i.e., $V_{AB} = 12$ V (for this circuit) and the derivation that substantiates the result. While we're discussing open circuits, we should point out that a short circuit is defined as an element whose voltage is zero for all time, no matter how much current is flowing through it.

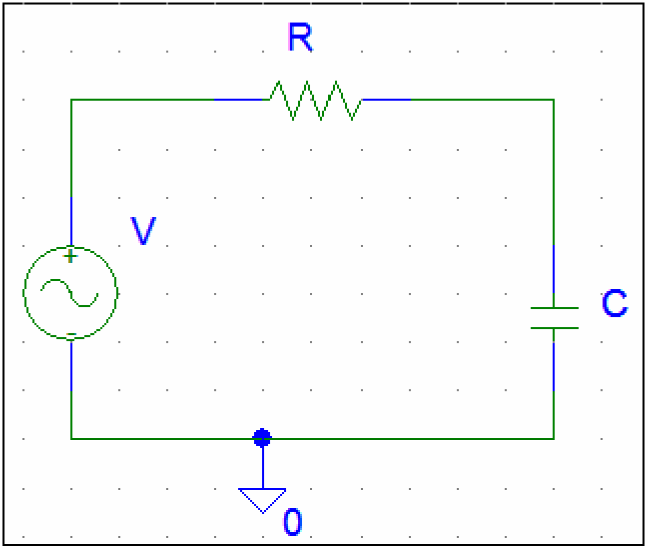
The capacitor voltage in the above RC circuit behaves the way it does because the capacitor can store charge. When the switch is first thrown, positive charges flow from the battery to the capacitor, producing a voltage across the capacitor. As time goes on, the capacitor voltage increases until it is equal to the battery voltage. Once the capacitor voltage is equal to the battery voltage, positive charges at the battery are repelled by the capacitor charge, and the current stops flowing. The voltage V across a capacitor is related to the charge Q it contains as

$$Q = C \cdot V$$

where C is the capacitance in farads. The derived unit "farad" then, has the fundamental unit coulombs per volt.

The fact that a fully charged capacitor behaves like an open circuit to a dc voltage explains, in part, why the capacitor can block the dc part of a composite signal (a signal that has both a dc part and an ac part).

To understand why the capacitor can pass some ac signals, we note that if the period T of an ac signal is less than the time constant τ, the ac signal can partially charge and discharge the capacitor on each cycle of the ac, and therefore "pass through the capacitor." Consider the circuit shown below.



We perform the following thought experiment: set the amplitude of the ac signal source to one volt, then sweep the frequency from zero to infinity while measuring the voltage across the capacitor and the current through the capacitor. The magnitude of the capacitor current is found to be

$$I_C = \frac{1}{\sqrt{R^2 + \frac{1}{(2\cdot\pi\cdot f\cdot C)^2}}}$$

The magnitude of the capacitor voltage is found to be

$$V_C = \frac{\frac{1}{(2\cdot\pi\cdot f\cdot C)}}{\sqrt{R^2 + \frac{1}{(2\cdot\pi\cdot f\cdot C)^2}}}$$

If we divide the capacitor voltage by the capacitor current (recall that $R = \frac{V}{I}$) we can define the ”resistance” of the capacitor to ac (sinusoidal) current flows as

$$\frac{V_C}{I_C}=\frac{\left(\dfrac{\dfrac{1}{(2\cdot\pi\cdot f\cdot C)}}{\sqrt{R^2+\dfrac{1}{(2\cdot\pi\cdot f\cdot C)^2}}}\right)}{\left(\dfrac{1}{\sqrt{R^2+\dfrac{1}{(2\cdot\pi\cdot f\cdot C)^2}}}\right)}=\frac{1}{(2\cdot\pi\cdot f\cdot C)}$$

The magnitude of the “resistance” of the capacitor to ac currents is called the capacitive reactance X_C. So the ac voltage of a capacitor is related to the ac current flowing through it as

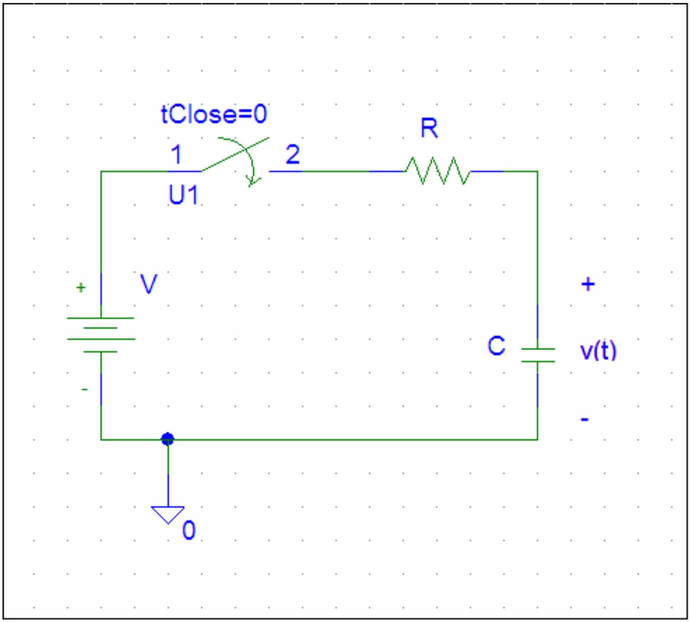
$$V_C = \left(\frac{1}{2 \cdot \pi \cdot f \cdot C} \right) \cdot I_C = X_C \cdot I_C$$

X_C has units of ohms. In our microphone-amplifier example, as long as $X_C << R_{in}$ the capacitive reactance can be ignored.

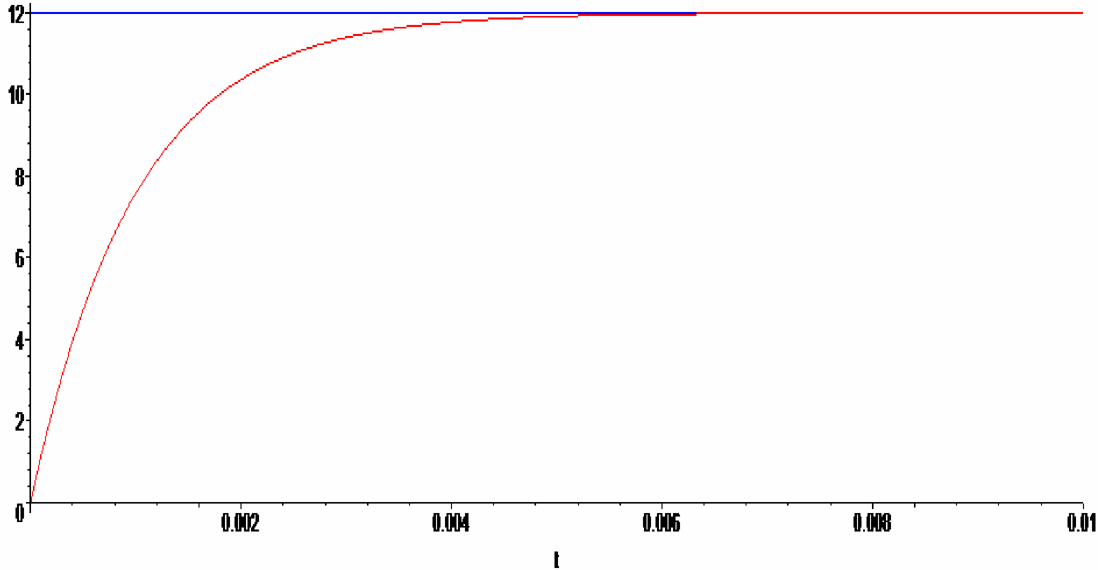
Since $R_{in} = 10,000\,\Omega$ and $X_C = \dfrac{1}{2 \cdot \pi \cdot 1000 \cdot 0.000001} \approx 160$, we may safely ignore the capacitive reactance and say that the capacitor allows the ac signal to pass through it while blocking the dc. In general when using a capacitor as a dc blocking capacitor it is necessary to choose a large enough capacitance so that $X_C < \dfrac{R}{10}$ where R is the smallest resistor in the circuit. It should be pointed out that when we change the frequency of the voltage source, we change not only the magnitudes (amplitudes) of the sinusoidal current and voltage of the capacitor but also their phase relationship with the source voltage. We'll have more to say about this in a subsequent chapter.

Signal Sources

We saw, in the last chapter, that a capacitor charges in a well defined time that depends on the values of R and C. In this chapter we exploit this property to construct a number of useful (and one fanciful) devices. If one wishes to test an amplifier, for example, it is necessary to provide an input signal to the amplifier. As we have already seen, if we build the following circuit



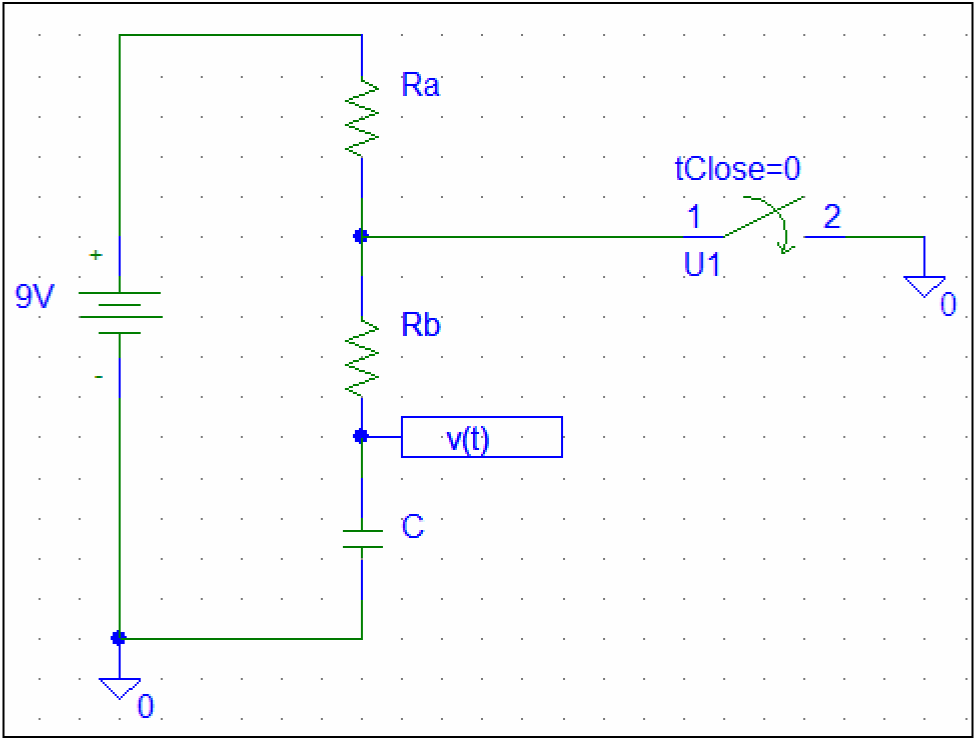
We obtain the output voltage v(t) as follows



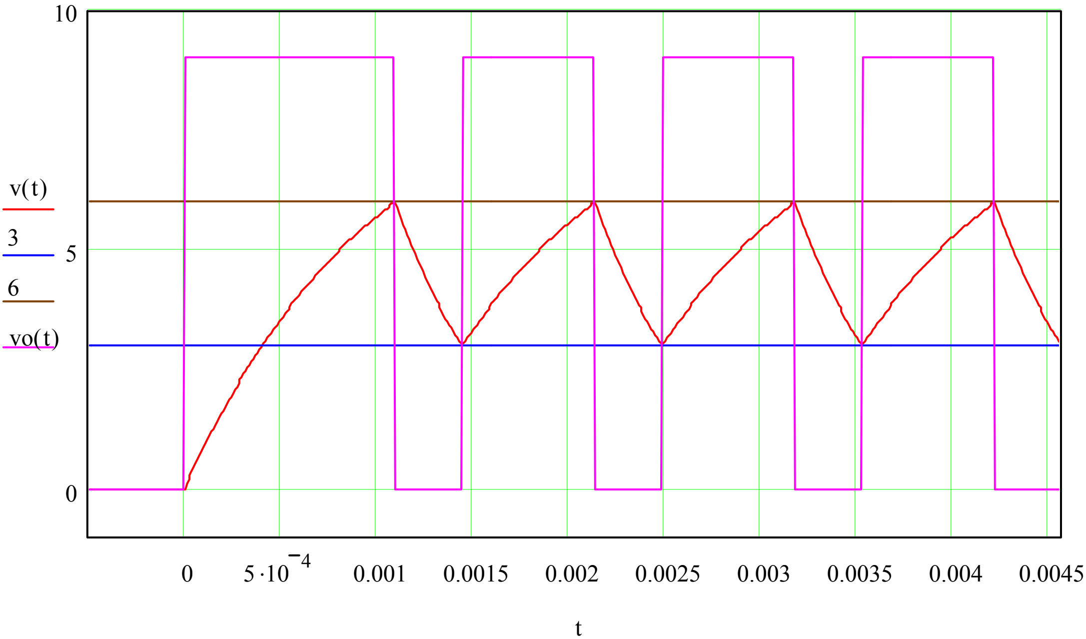
Where v(t) is described by

$$v(t)=V\cdot\left(1-e^{-\frac{t}{R\cdot C}}\right)$$

By alternately charging and discharging the capacitor we can produce a time varying voltage. Consider the circuit shown below.

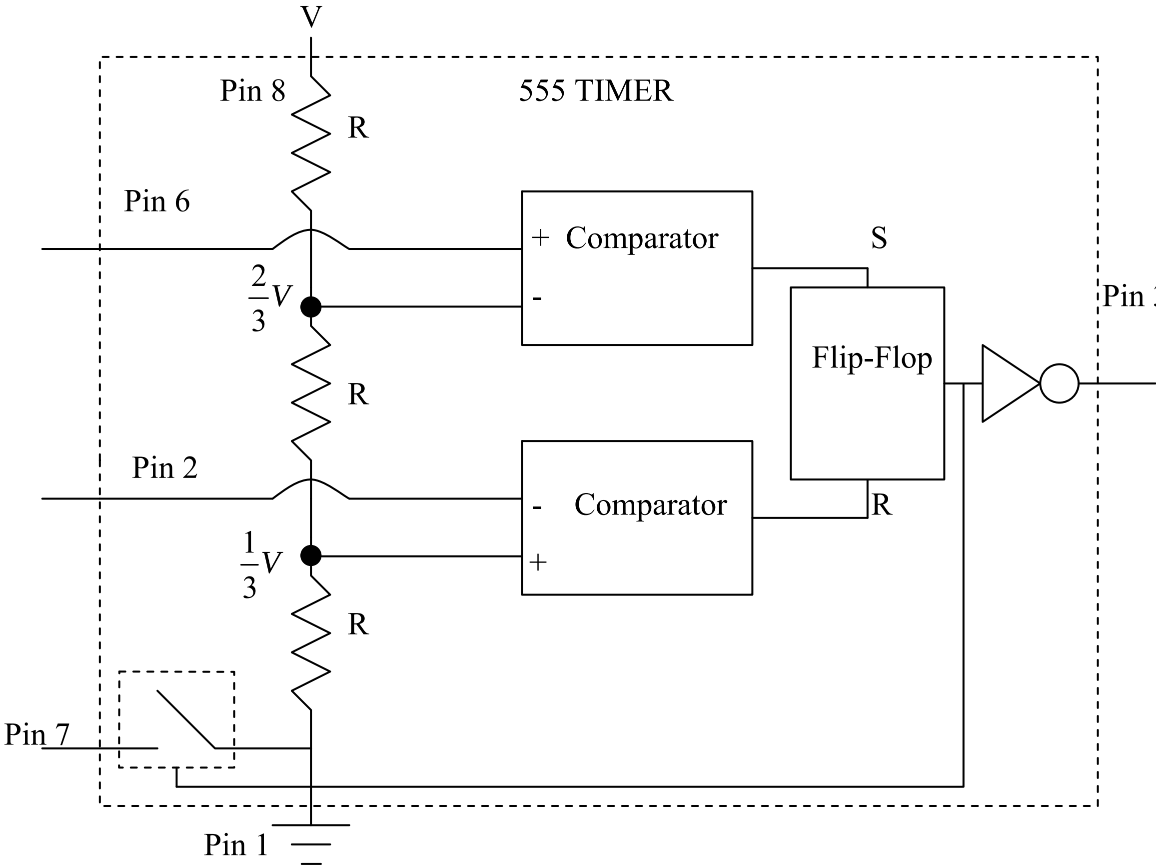


If we let the capacitor charge until v(t) = 6V, then close the switch to discharge the capacitor until v(t) = 3V, then open the switch and let the capacitor charge again, and continue this pattern we would get a voltage like the one shown below.

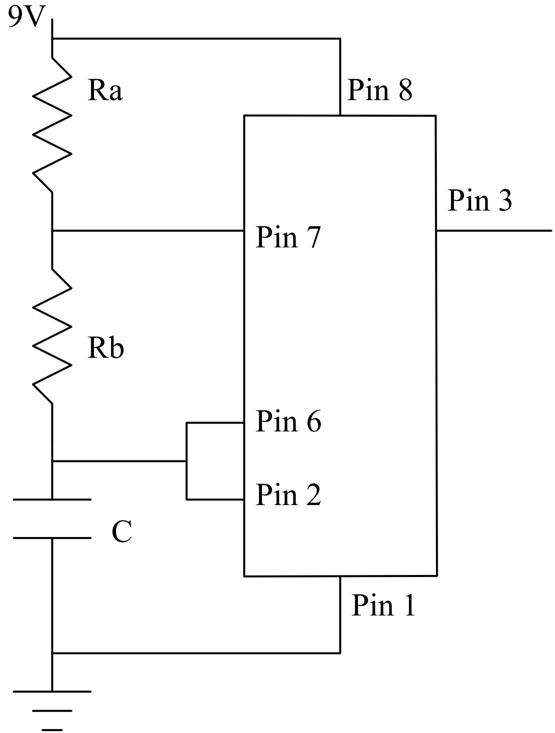


The device that handles this is called the NEC555 timer (or LM555 timer, or simply the 555 timer). NEC was one of the early developers of the 555 timer and 555 was the part number assigned to it.

The 555 timer contains the following circuitry shown in block diagram form.



To produce a voltage like vo(t) shown above we connect the 555 into the following circuit: the circuit is called a free-running or astable multi-vibrator.



555 Astable Multi-vibrator

We can calculate the frequency of the output as follows: ta is the time where the rising waveform passes through $\frac{1}{3}\cdot V$, tb is the time when the waveform reaches $\frac{2}{3}\cdot V$ and tc is the time when the descending waveform reaches $\frac{1}{3}\cdot V$

$$V\cdot\left(1-e^{-\frac{ta}{(Ra+Rb)\cdot C}}\right)=\frac{1}{3}\cdot V$$ so

$$e^{\frac{-ta}{(Ra+Rb)\cdot C}}=\frac{2}{3}$$ Then taking the ln (natural log) of both sides we have

$$\frac{-ta}{(Ra+Rb)\cdot C}=\ln\left(\frac{2}{3}\right)$$ which gives

$$ta = \left(\ln(3) - \ln(2)\right) \cdot (Ra + Rb) \cdot C$$

From

$$V \cdot \left(1 - e^{-\frac{tb}{\left(Ra + Rb\right) \cdot C}}\right) = \frac{2}{3} \cdot V$$ we obtain

$$e^{\frac{-tb}{\left(Ra + Rb\right) \cdot C}} = \frac{1}{3}$$ which gives

$$tb = \ln(3) \cdot \left(Ra + Rb\right) \cdot C$$

So

$$tb - ta = \ln(2) \cdot \left(Ra + Rb\right) \cdot C$$

This is the time the capacitor is charging from 3 V to 6 V.

A similar derivation starting with

$$\frac{2}{3}\cdot V\cdot e^{-\frac{tc-tb}{Rb\cdot C}}=\frac{1}{3}\cdot V$$

gives

$$tc-tb=\ln(2)\cdot Rb\cdot C$$

The time for a complete cycle is, then

$$T=(tb-ta)+(tc-tb)=\ln(2)\cdot(Ra+2\cdot Rb)\cdot C$$

The frequency f is

$$f=\frac{1}{T}=\frac{\frac{1}{\ln(2)}}{(R1+2\cdot Rb)\cdot C}\approx\frac{1.44}{(R1+2\cdot Rb)\cdot C}$$

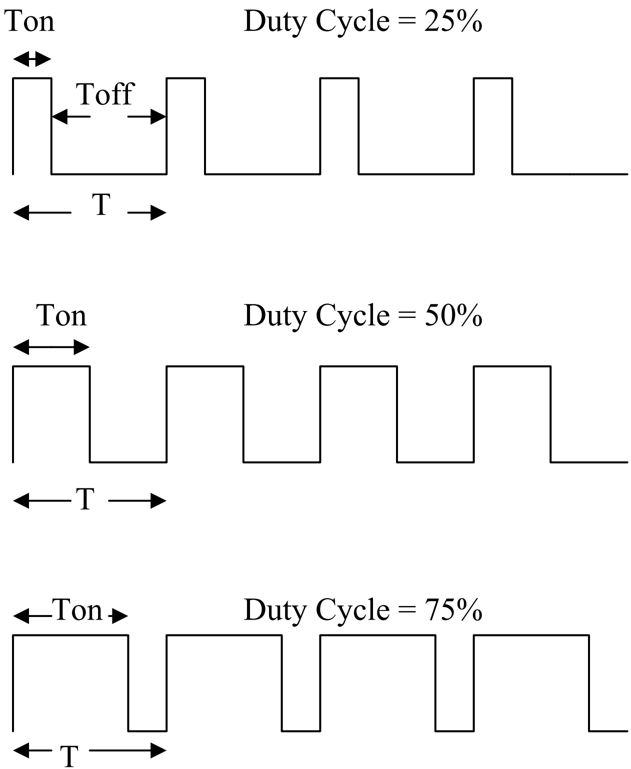
Before we modify the circuit to produce a useful signal with a 50% duty cycle, we pause to consider a toy version of a musical instrument called a Theremin. The Theremin is played without touching the instrument; the instrument has two antennae which the player uses to control the instrument. As the player moves his or her hand toward or away from the pitch antenna, the pitch rises and falls. As the player moves his or her other hand toward or away from the loudness antenna, the volume of the sound rises and falls. Our toy version has only pitch control capabilities.

We can change the pitch if we can change the value of either Ra or Rb as we move a hand closer to, or away from the circuit. We exploit the property of cadmium sulfide that causes its resistance to change in the presence of light. The component is called a light dependent resistor (LDR), and consists of a piece of cadmium sulfide with a small window. Our LDR has a resistance of about 5k ohms in the dark and 200 – 300 ohms in a well-lit room. We will use the LDR as Rb, and we'll have an oscillator whose frequency depends on how much light is falling on the face of the LDR.

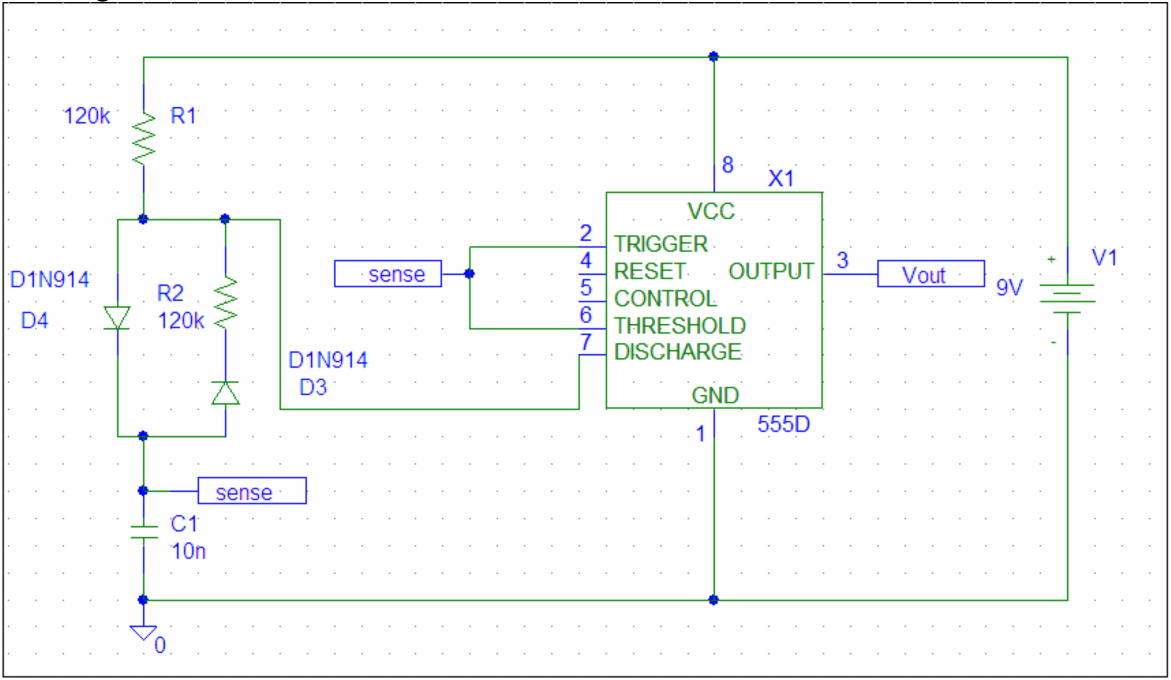
To get a symmetric waveform with a 50% duty cycle, we first pause to define duty cycle. The duty cycle of a waveform is the ratio of its on-time to its total time expressed as a percentage.

$$D = \frac{Ton}{T} \cdot 100$$

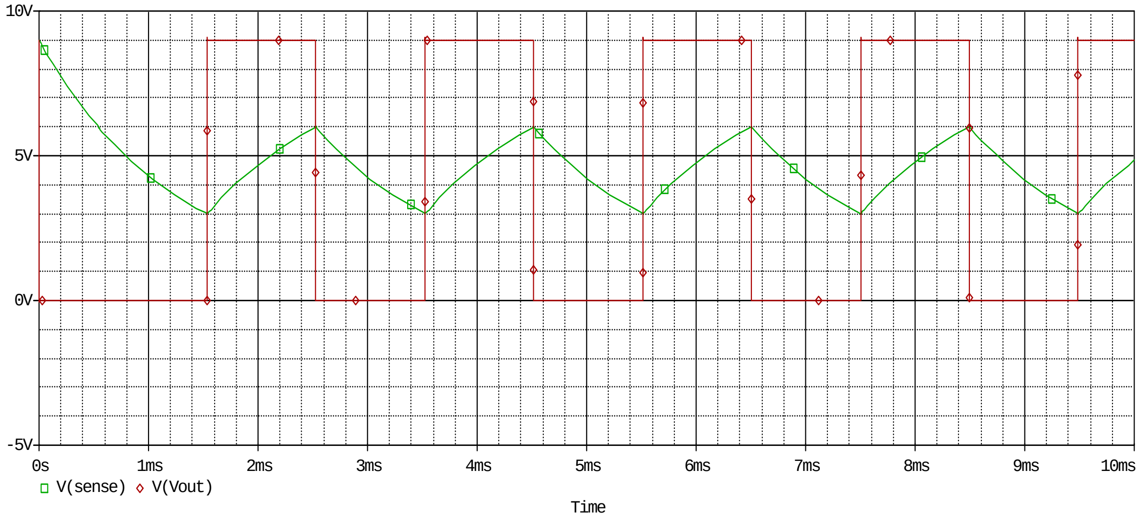
This concept is most easily understood by a few examples.


Ton
Duty Cycle = 25%
Toff
T
Ton
Duty Cycle = 50%
T
Ton
Duty Cycle = 75%
T

We obtain symmetric charge and discharge times by steering the charge current and discharge currents with diodes as shown below.



The capacitor and output waveforms are shown below: note that their duty cycles are 50%.



The approximate output frequency is

$$f=\frac{1}{T}=\frac{\frac{1}{\ln(2)}}{(R1+Rb)\cdot C}\approx\frac{1.44}{(R1+Rb)\cdot C}$$

In The Lab

Build the Theremin from the figure labeled 555 Astable Multivibrator using the 5k LDR as Rb, the 4.7k as Ra and the 100 nF capacitor (marked 104) as C.

Capacitor markings

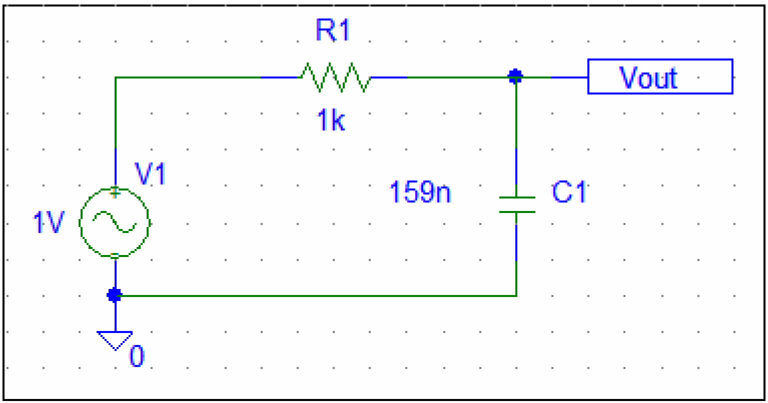
There are, broadly speaking two kinds of capacitors: electrolytic and non-polarized. Capacitors whose value is greater than $1\mu F$ are usually electrolytic, and capacitors whose value is less than $1\mu F$ are usually non-polarized. (There are several different kinds of non-polarized capacitors with different characteristics, but they are all marked the same way so we shall consider them together.) Electrolytic capacitors (and some non-polarized capacitors are marked with their capacitances in micro-farads: $1\mu F$, $10\mu F$, $100\mu F$, $0.1\mu F$, $0.01\mu F$, $0.001\mu F$, etc. Most non-polarized capacitors are marked with a series of three digits that are interpreted in the same way as the bands on resistors are interpreted. That is, the first two digits are the value and the last digit is the multiplier. So 104 would be 10 times $10^4 = 10^5$ **pico-farads**. Since one pico-farad $=10^{-12}$ Farads,

$$104 -> 10^5 \cdot 10^{-12} F = 10^{-7} F = 0.1 \cdot 10^{-6} F = 0.1\mu F = 100 nF$$

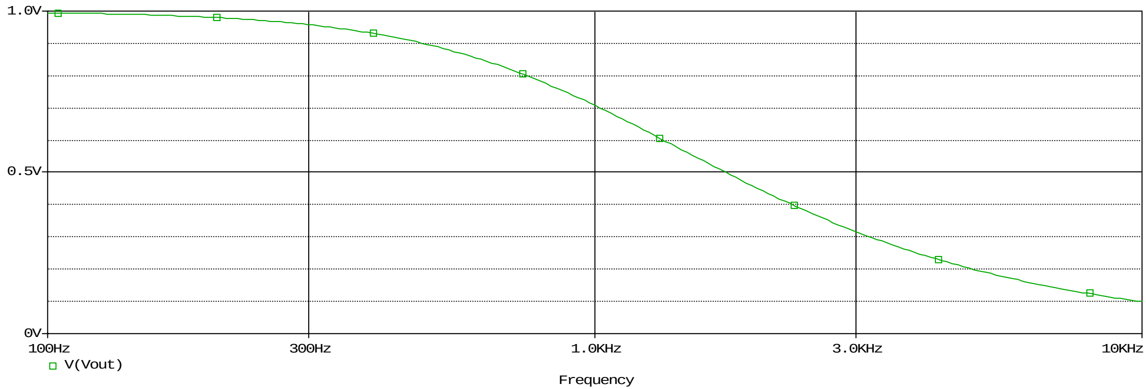
In the last chapter we noted that the capacitor reactance (its ac "resistance") is frequency dependent as

$$X_C = \frac{1}{2 \cdot \pi \cdot f \cdot C}$$

We can exploit the fact that the capacitor reactance is frequency dependent, and the fact that a square wave can be described by Fourrier's theorem as a sum of sine waves to change the square wave output from our 555 into a sine wave with the same period. The circuit used to accomplish this is called a filter.
Consider the circuit below, where the frequency of the source is changed from 100 Hz to 1000 Hz.



The result of plotting the output voltage vs. frequency is show below.



Note that the frequency is logarithmic, i.e., we plot against log (f) rather than f. This is standard practice when the range of frequencies is large. It is also common practice to convert voltages that have magnitudes that vary over a wide range to decibels.

The Bel (after Alexander Graham Bell) was invented by telephone engineers trying to describe the power loss over very long cables: these powers varied over several orders of magnitude. The Bel describes a ratio of powers and is defined as

$$P_{Bel} = \log\left(\frac{P_{out_{Watts}}}{P_{in_{Watts}}}\right)$$

Since the Bel is an inconveniently large unit the standard unit is the deci-Bel or decibel, defined as

$$P_{deci-Bel} = 10 \cdot \log\left(\frac{P_{out_{Watts}}}{P_{in_{Watts}}}\right)$$

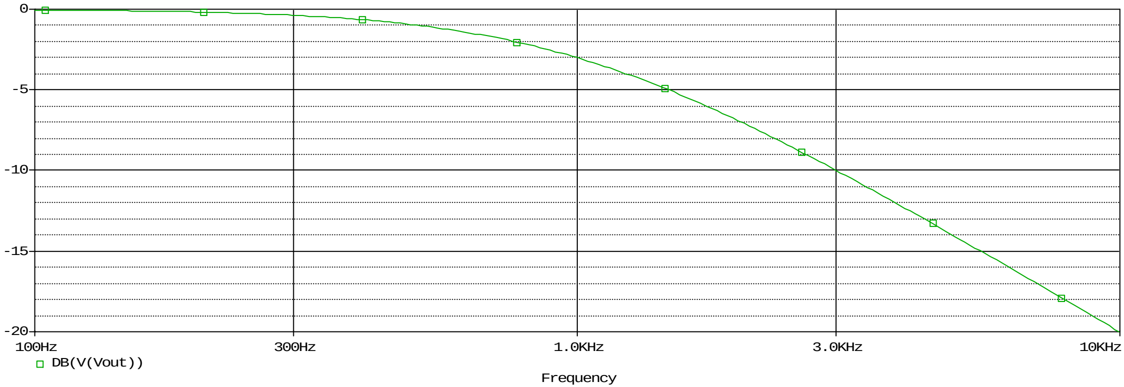
To take advantage of the decibel for voltages and current we must recall that P = V * I and V = R * I. Then P = I * I * R or P = I^2 *R or P = V * V / R or P = V^2 / R. Making the substitution for power in terms of V we have

$$V_{deci-Bel} = 10 \cdot \log\left(\frac{\frac{V^2_{out}}{R}}{\frac{V^2_{in}}{R}}\right) = 10 \cdot \log\left(\frac{V^2_{out}}{V^2_{in}}\right) = 20 \cdot \log\left(\frac{V_{out}}{V_{in}}\right)$$

A similar development for I gives

$$I_{deci-Bel} = 10 \cdot \log\left(\frac{I^2_{out} \cdot R}{I^2_{in} \cdot R}\right) = 10 \cdot \log\left(\frac{I^2_{out}}{I^2_{in}}\right) = 20 \cdot \log\left(\frac{I_{out}}{I_{in}}\right)$$

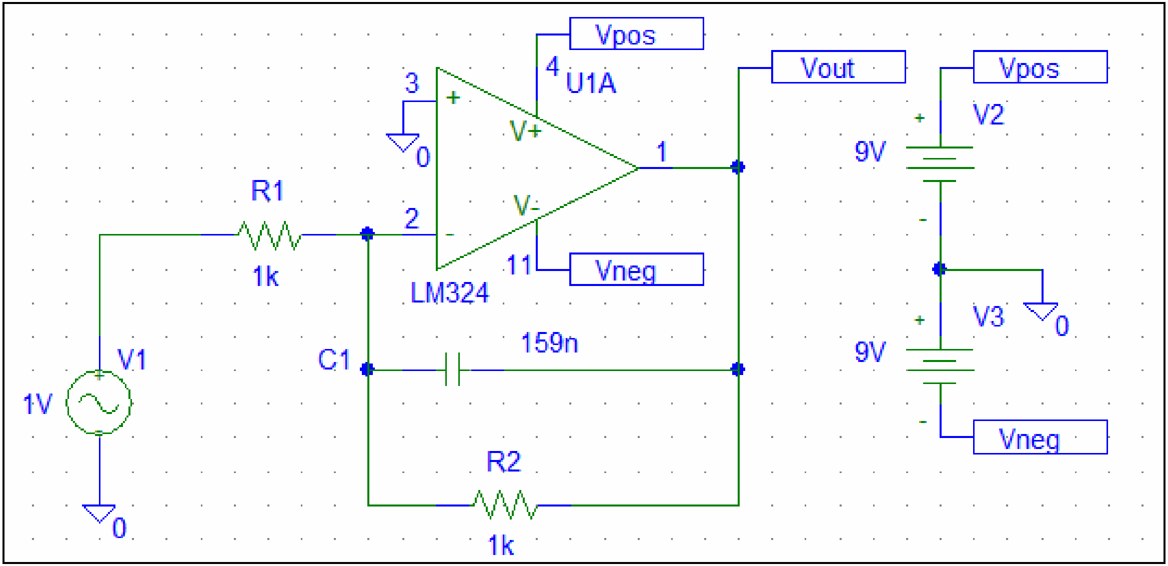
Plotting the ratio of the output voltage to the input voltage of the above circuit in decibels gives:



This filter response is called a low-pass response for obvious reasons.

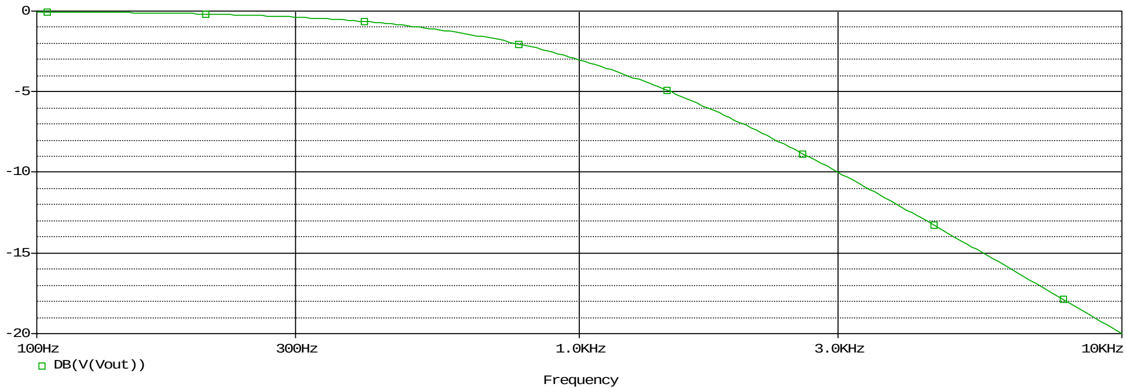
Our plan is to use a low-pass filter to filter out all of the harmonic components (sine waves with frequency 2*f, 3*f, 4*f, etc) leaving only the fundamental (the sine wave whose frequency is f). To do this we must improve our filter to roll off more steeply at high frequencies. We take a necessary first step by introducing an op-amp. Our first op-amp circuit has the same roll off characteristics as the simple R C filter, but is easy to modify for steeper roll off.

The first order (there is one capacitor) active filter (uses an op-amp) is shown below.



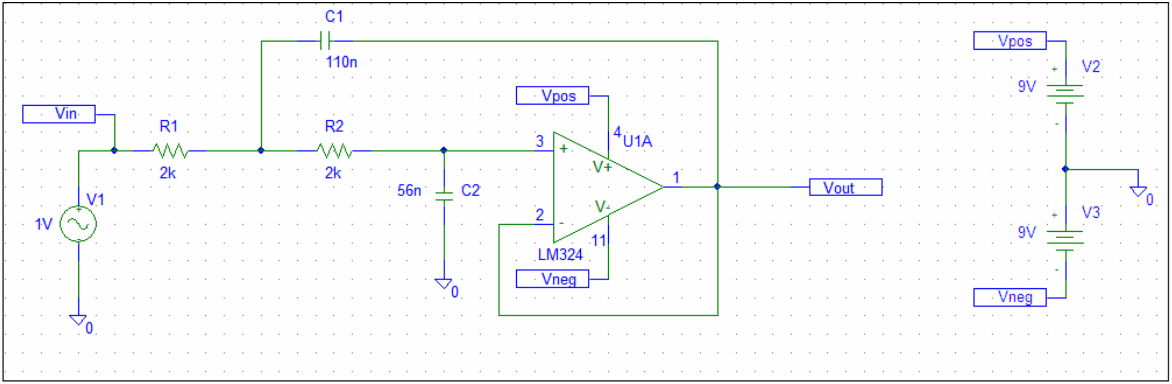
First-order Low-pass Filter

Its frequency response is identical to that of the passive RC filter and is shown below.

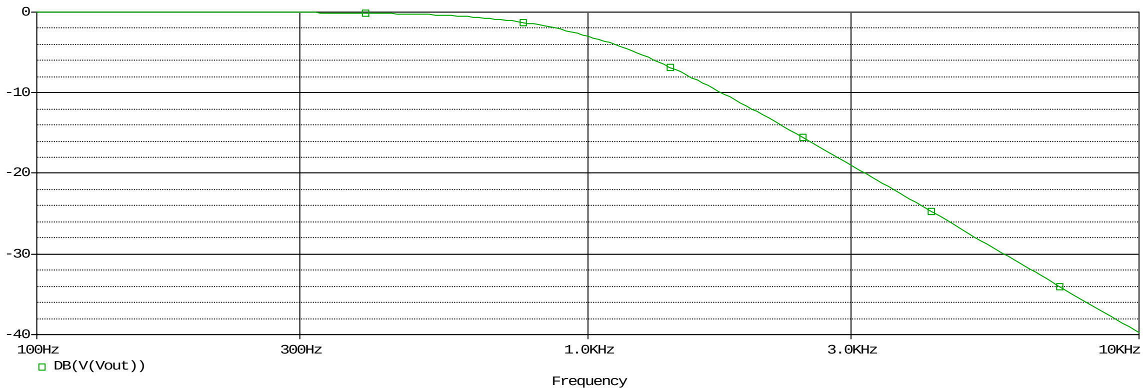


First-order Low-pass Filter Response

The real advantage of the op-amp version is that its frequency roll off is easily improved. Consider the second order low-pass filter shown below.

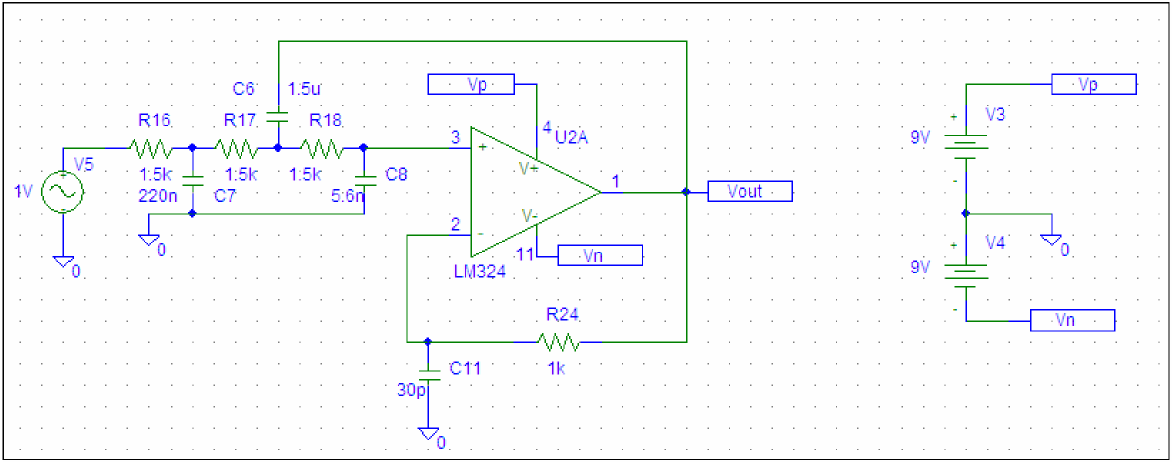


Second-order Low-pass Filter

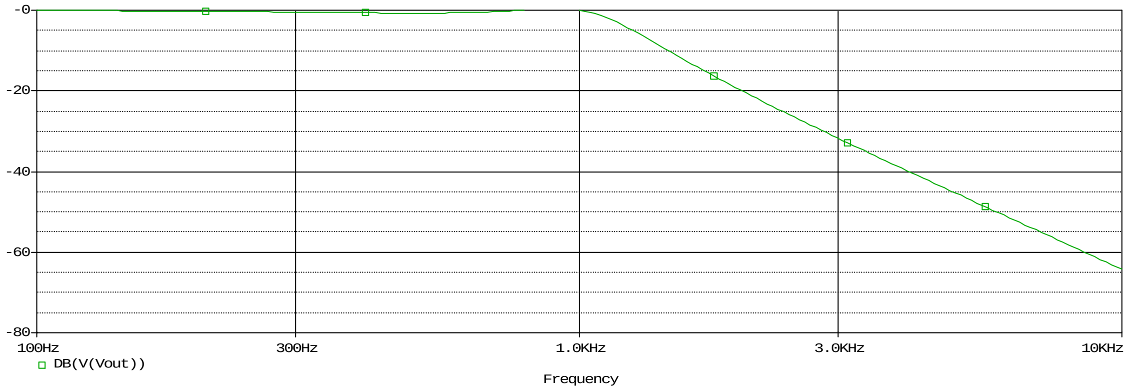


Second-order Low-pass Filter Response

Notice that the second order filter has a steeper drop-off in frequency than the first order filter does. By adding a third capacitor we form the third-order filter shown below.

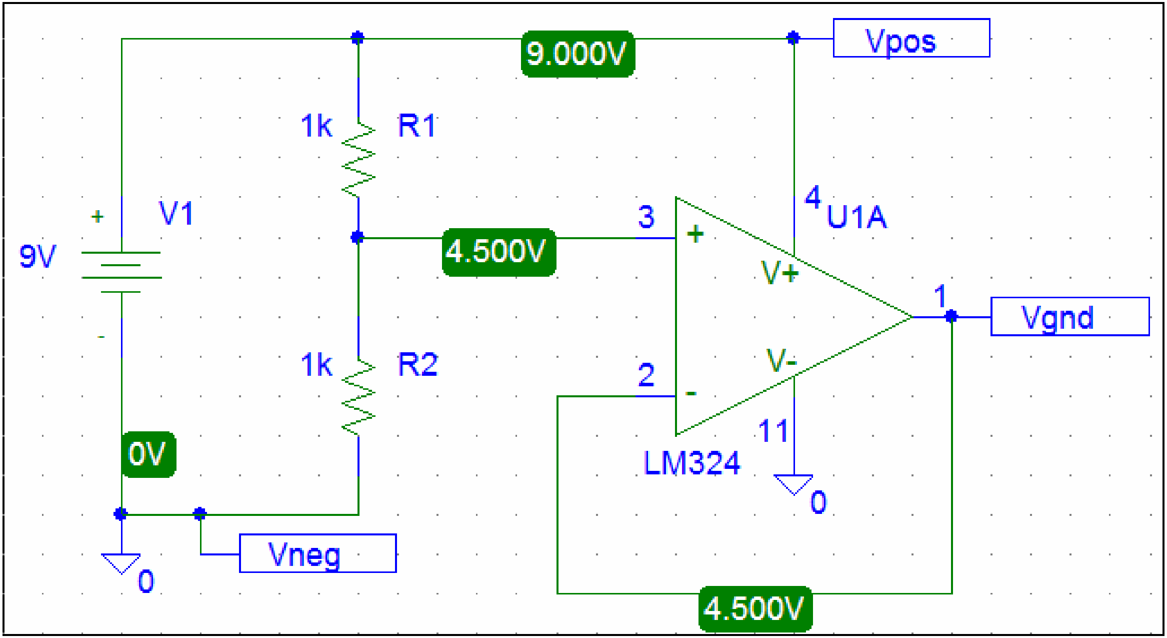


Third-order Low-pass Filter

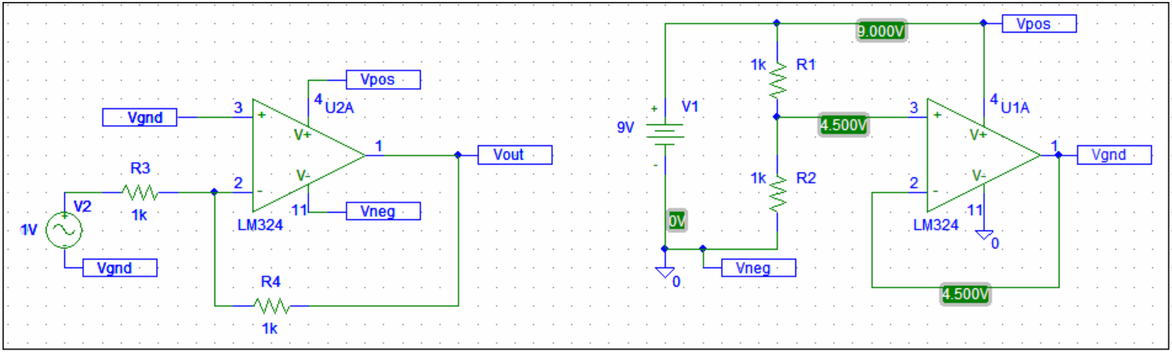


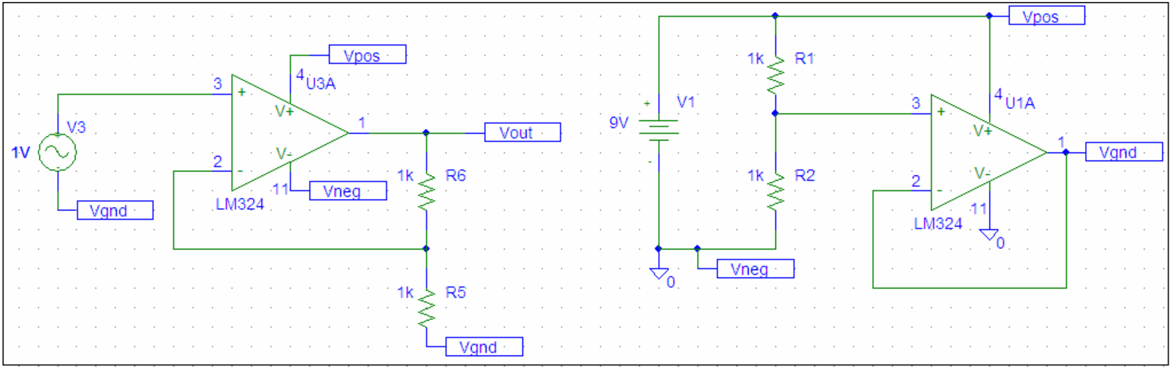
Third-order Low-pass Filter Response

Our intention is to connect the output from a 555 timer to a third-order low-pass filter to produce a sine wave output. Since out 555 timer is powered by a single 9V source, we seek a way to run our op-amp circuits from a single 9V source. We start by noticing that there is no ground pin on an op-amp: the op-amp *infers* ground as the voltage half-way between its positive supply voltage and its negative supply voltage. If we connect the op-amp positive supply pin to +9V and its negative supply pin to ground, we must produce a voltage half-way between 0V and 9V (i.e. 4.5V) to use as a pseudo-ground. This point must remain at 4.5V no matter how much current is drawn from the 4.5V supply. Consider the voltage divider shown below.



The op-amp is necessary to keep Vgnd from changing when current is drawn from Vgnd. We can then modify our inverting and non-inverting amplifiers as follows.

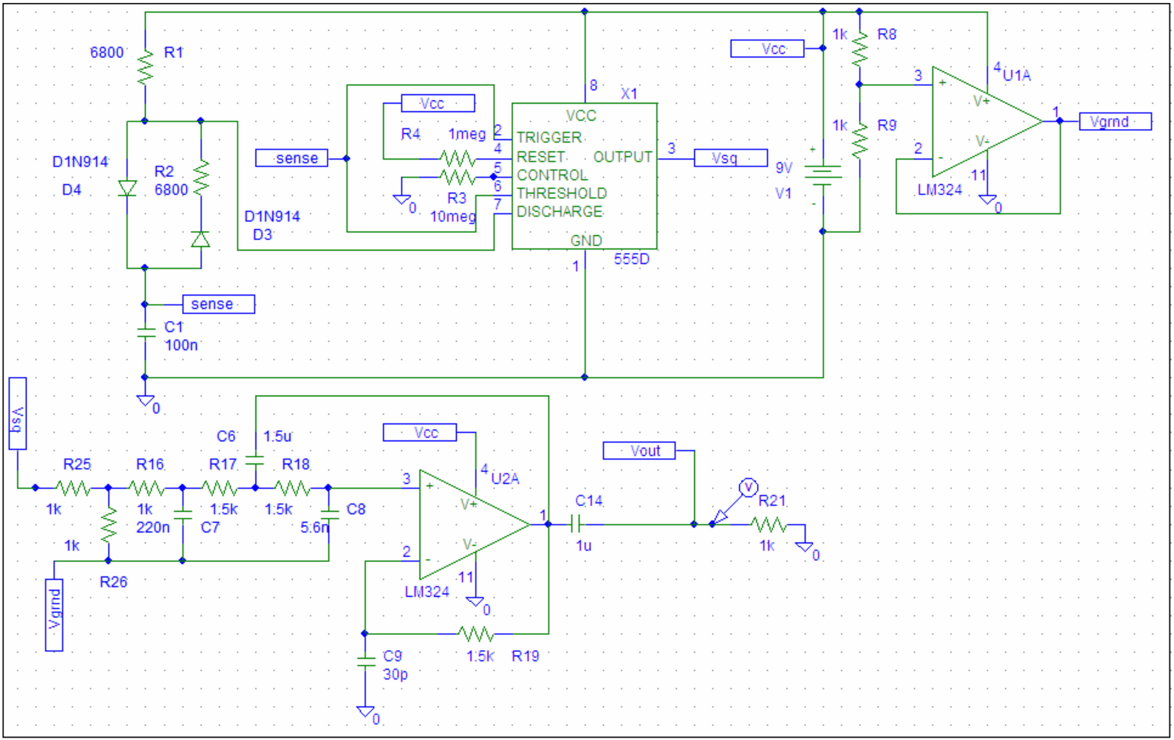




Note that all previous references to ground now go to Vgnd.

In the Lab

Construct the following circuit and measure and print the waveforms at Vsq and Vout with the oscilloscope.



Most of us are aware that when we bring a ferro-magnetic material (iron or steel, for example) near to a magnet, that the magnet exerts an attractive force on the ferro-magnetic material, and that the magnitude of the force grows weaker with distance. (The distance between the magnet and the ferro-magnetic material.) Most of us are also aware that the two ends of a magnet behave differently: that there is a polarity associated with the ends of a magnet. Suppose we have two identical magnets as shown below in Fig. 1.

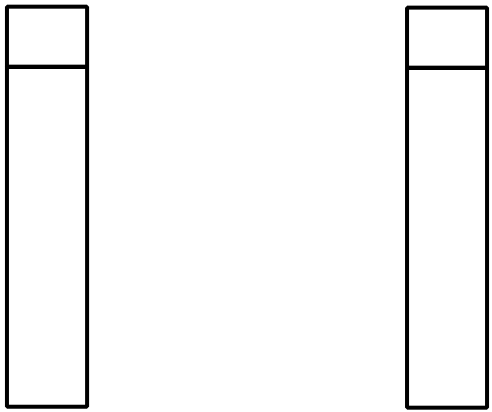


Figure 1 Two identical magnets

If the magnets are brought together as shown in Fig. 2, the force between them is as shown in Fig. 2, where we see that like ends repel each other and unlike ends attract each other.

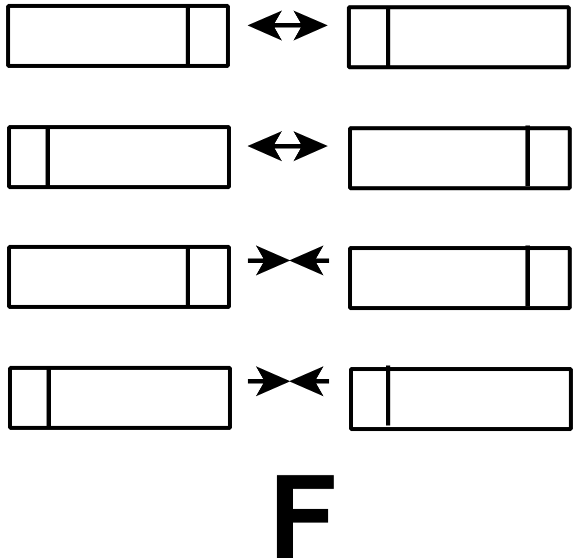


Figure 2 Polarity of Magnetic Forces

In about 100 B.C. Chinese scientists noticed that if a magnetized needle is carefully balanced on a pivot so that it is free to rotate, one end of the needle points approximately toward the geographic north pole. (The same effect can be produced by floating a magnetized needle on a small bit of cork.) Thus the marked end of the magnets in Fig. 2 are called north magnetic poles and the unmarked ends are called south magnetic poles. This phenomenon occurs because the earth has a magnetic field that is approximately aligned with the earth's geographic poles. It is interesting to note that because we decided to call the end of the magnet that points to the north geographic pole the north magnetic pole, that the earth has a south magnetic pole at the north geographic pole. Unfortunately the work of the Chinese was lost for many centuries and it was not until the 1200s A.D. that the compass was used regularly for ship navigation by the Europeans.

Suppose we pass currents through pair of parallel wires as shown in Fig. 3 below. (Note that these currents flow in the same direction.) When the currents are flowing we find that there is an attractive force between the wires as shown. (Of course if we bring a piece of ferro-magnetic material near either of the conducting wires it would be attracted to the current-carrying wire.)

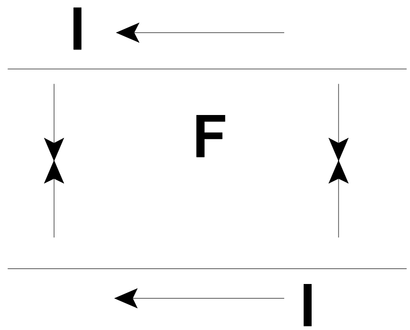


Figure 3 Attracting Wires

If we reverse one of the currents as shown in Fig. 4 below , (so that the currents are flowing in opposite directions) we find that the force is repulsive. (The wires tend to move away from each other.)

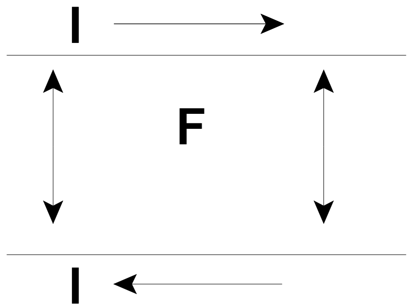


Figure 4 Repelling wires

If we investigate the magnetic force caused by one current-carrying conductor by itself we find that the force is everywhere perpendicular to the direction of the current flow. This force is fairly weak, so we use a compass needle to determine the direction of the force, since a compass needle aligns itself with the closest, strongest magnetic field in its vicinity. The results are illustrated in Fig. 5 below. The current is flowing into the page, and the force can be seen to be tangent to circles drawn around the wire.

We can posit the “right-hand rule” for the direction of a magnetic field of a current carrying conductor as follows: “If the thumb of the right hand points in the direction of the current, then the magnetic field points in the direction of the curl of the fingers of the right hand.” In the study of electro-magnetic phenomena you will encounter other “right-hand rules.”

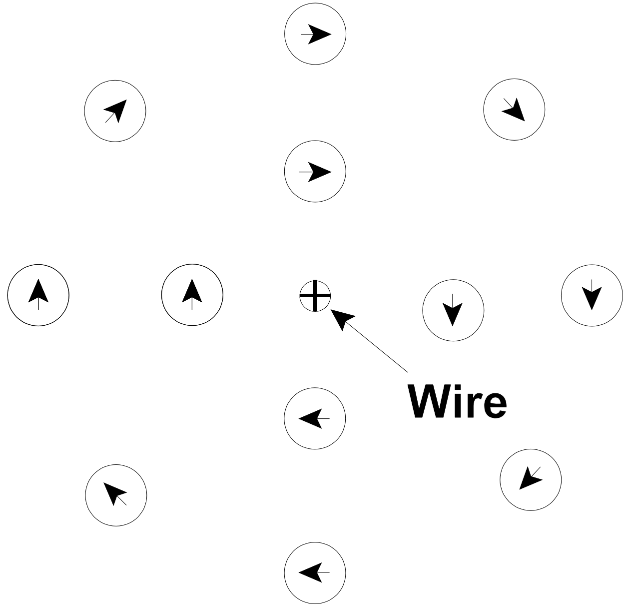


Figure 5 Magnetic Field Around a Current-Carrying Conductor

Note that the force is tangential to concentric circles drawn around the wire (which is carrying current into the page i.e. away from the reader).

We can use the results of Fig. 5 to explain the results of Figs. 3 and 4 by noting the attraction or repulsion of differential lengths of the magnetic field lines. Note in Fig. 6 the use of the **standard convention** that a dot means the current is flowing out of the page toward the reader, (the tip of an arrow coming toward the reader), and a cross means the current flowing into the page away from the reader (the feathered end of an arrow traveling away from the reader).

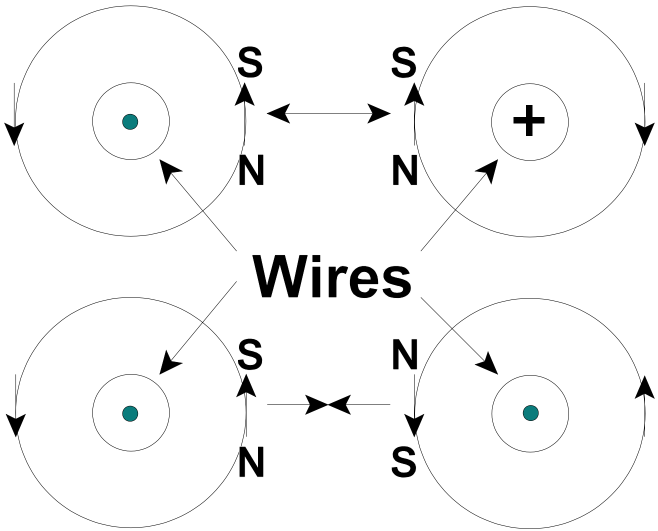


Figure 6 Currents in the Same Direction Attract
Currents in the Opposite Direction Repel

In Fig. 6, the wires are shown in cross-section with currents flowing according to the standard convention.

Another well known experiment involves wrapping a number of turns of copper wire around a cylindrical form as shown in Fig. 7 below. When an electrical current is passed through the wire, the entire assembly (called a solenoid) behaves exactly like a bar magnet, since the fields of the individual turns aid one another. (We will return to the direction of the magnetic force of a solenoid presently, for the moment we note that the behavior of the solenoid is indistinguishable from the behavior of a bar magnet.)

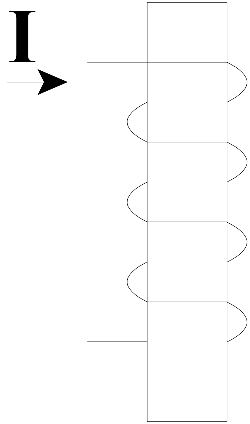


Figure 7 Electromagnet (Solenoid)

In fact, all magnetic effects are due to moving charges - in permanent magnets, electrons orbiting their nuclei are responsible for the magnetic force. Most students are well aware that increasing the current in the wire increases the strength of the magnetic force. In fact, the magnetic force is directly proportional to the current intensity (and to the number of turns). It is possible to use the electromagnet to characterize permanent bar magnets as follows. As shown in Fig. 8, find the heaviest mass that the bar magnet can hold. (This is done by increasing the mass m until the mass drops off the magnet, then slightly reducing the mass until the magnet can just hold it up.) The mass is then moved to the electro-magnet and the current is increased until the electro-magnet can just hold the mass up. The magnetic field intensity of the electro-magnet (and therefore of the bar magnet), is then given as

$$H = \frac{N\,I}{\ell} \quad \frac{Ampere\ Turns}{Meter}$$

where N is the number of turns, I is the current in Amperes, and ℓ is the length of the solenoid in meters.

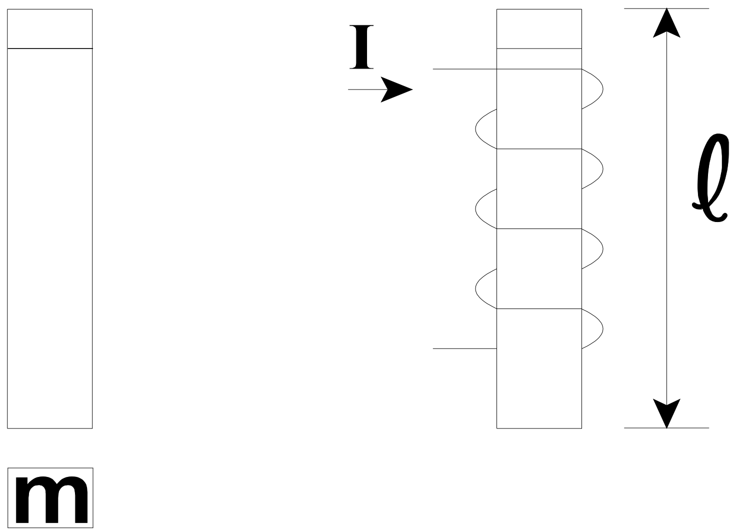


Figure 8 Characterization of Bar Magnet with an Electro-magnet

Two additional experiments help clarify our thinking about magnets. The first involves bringing the magnets together end to middle as shown in Fig. 9 below. Notice that there is no force generated left-to-right: only up and down.

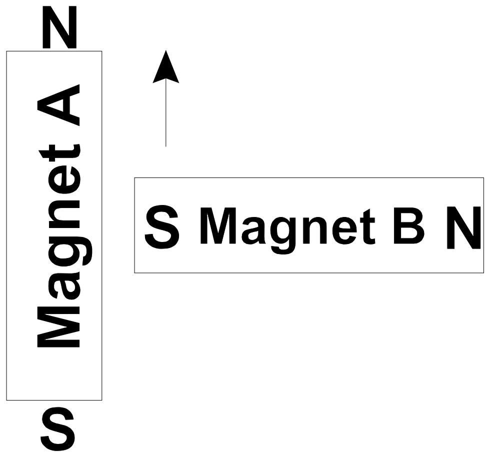


Figure 9 Magnets End to Middle

The other common experiment is to place a piece of paper over a magnet and sprinkle powdered iron on the paper. The well-known pattern shown below in Fig. 10 results. Note the complexity of these lines as opposed to the lines of force produced by charges (or gravity). The idea is that the magnetic force operates along the direction of these "lines of force," and that the magnitude of the force is proportional to the number of lines that are interacting, i.e. , more lines = more force. (By convention lies of force leave a north magnetic pole and enter a south magnetic pole.)

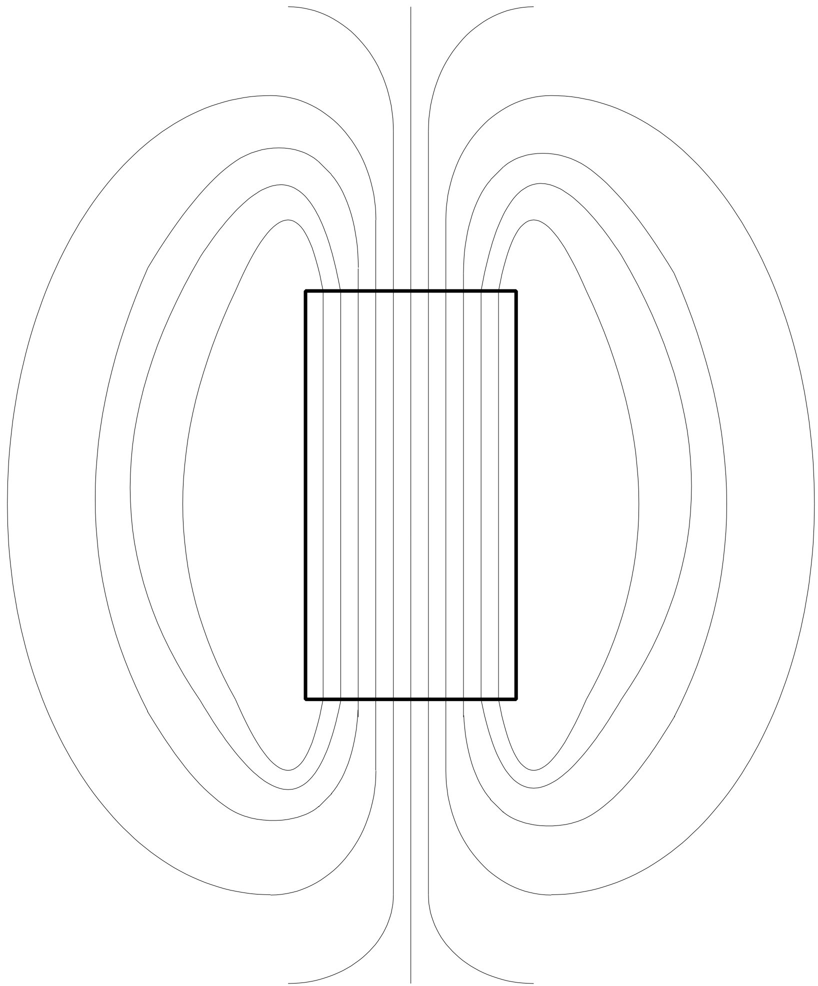


Figure 10 Magnetic Lines of Force

If the experiment is repeated with lots of tiny compasses instead of iron filings, the compass needles would point as in Fig. 11. Note that the lines (needles) point **away** from a north pole and **toward** a south pole.

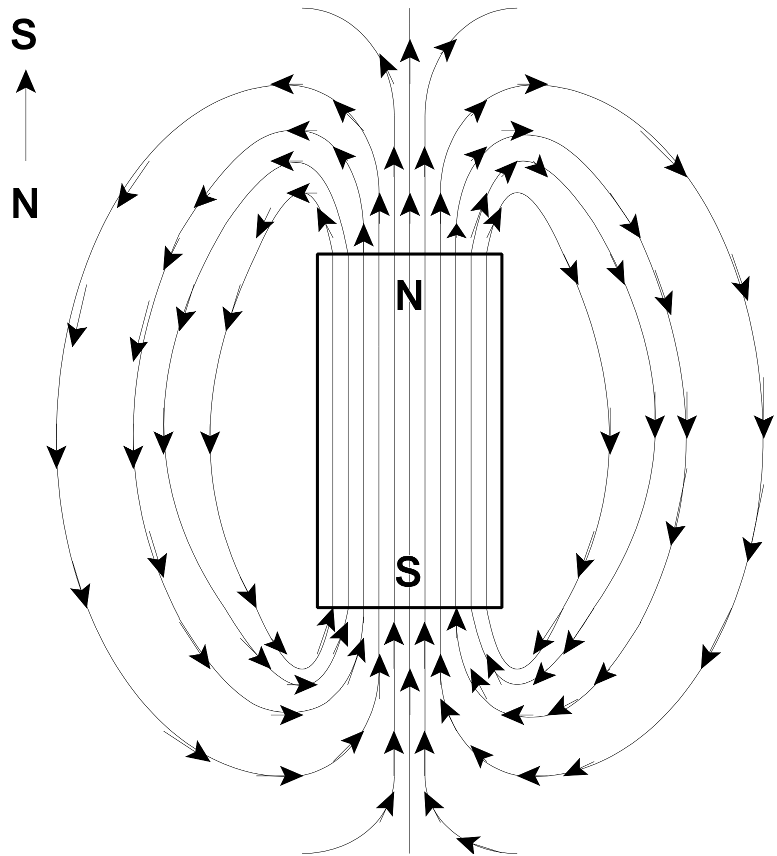


Figure 11 Magnetic Lines of Force with Compasses

Note the concentration of lines at the poles in Fig. 11.

Gravitational (or charge) “lines of force,” as shown in Fig. 12 below, are independent of the medium in which they exist. (It does not matter, for instance, whether an atmosphere is present or not, or whether there are any clouds or not - the gravitational force is unaffected.)

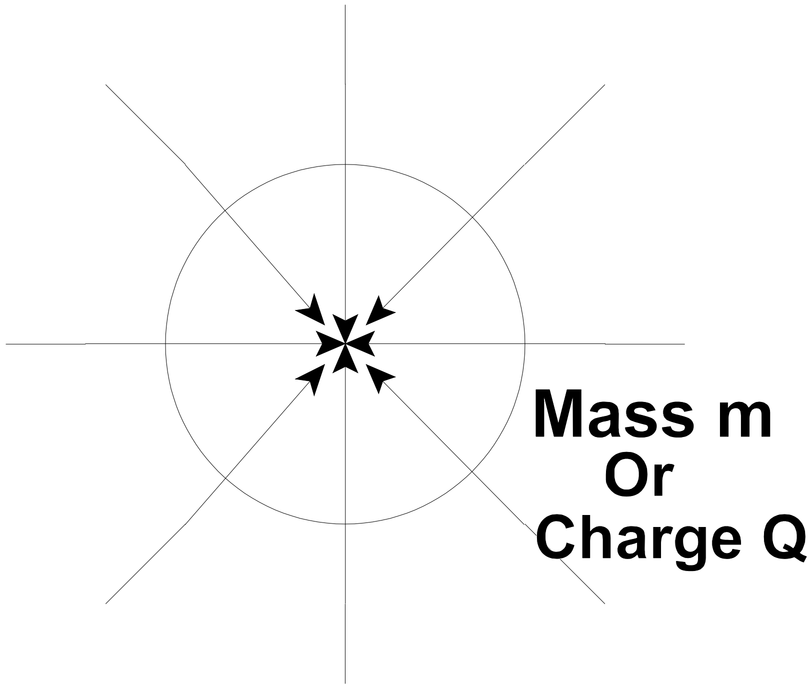


Figure 12 Gravitational (or Charge) Lines of Force

Another experiment that is not so well known, but which is very important to understanding magnetic forces, is illustrated in Fig. 13.

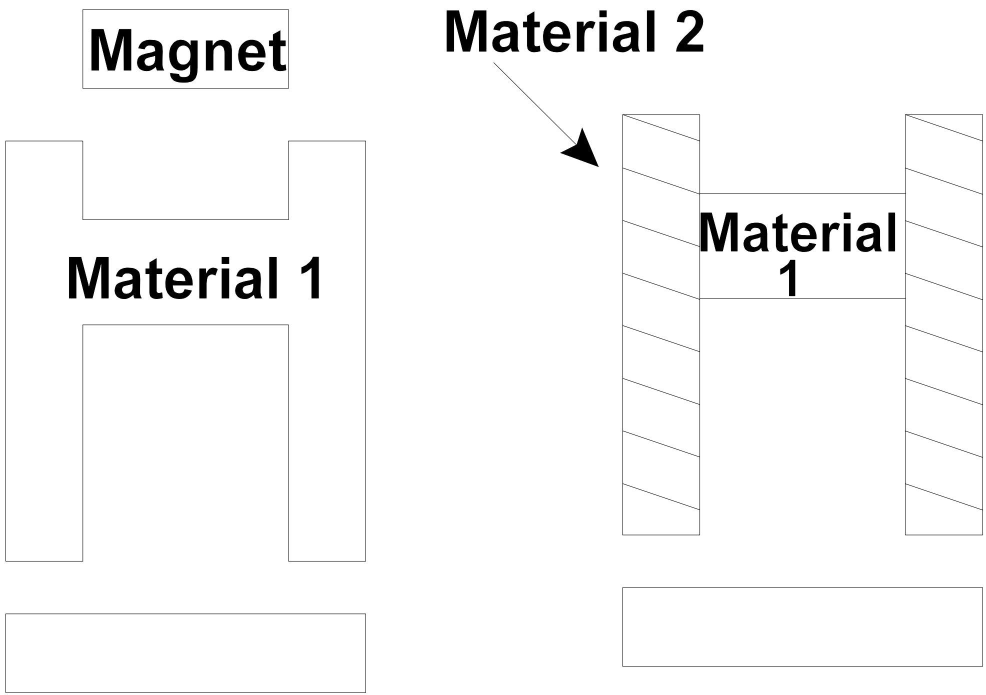


Figure 13 Magnet Holder Experiment

If the magnet is inserted into the holder on the left, the weight of the object that can be lifted is very small. If the same magnet is inserted into the holder on the right, the amount of weight that can be lifted is orders of magnitude larger. The explanation is that material 2 can bend the lines of force, whereas material 1 cannot. We usually say that the lines of force travel through material 1 more easily than they flow through the air or through material 2.

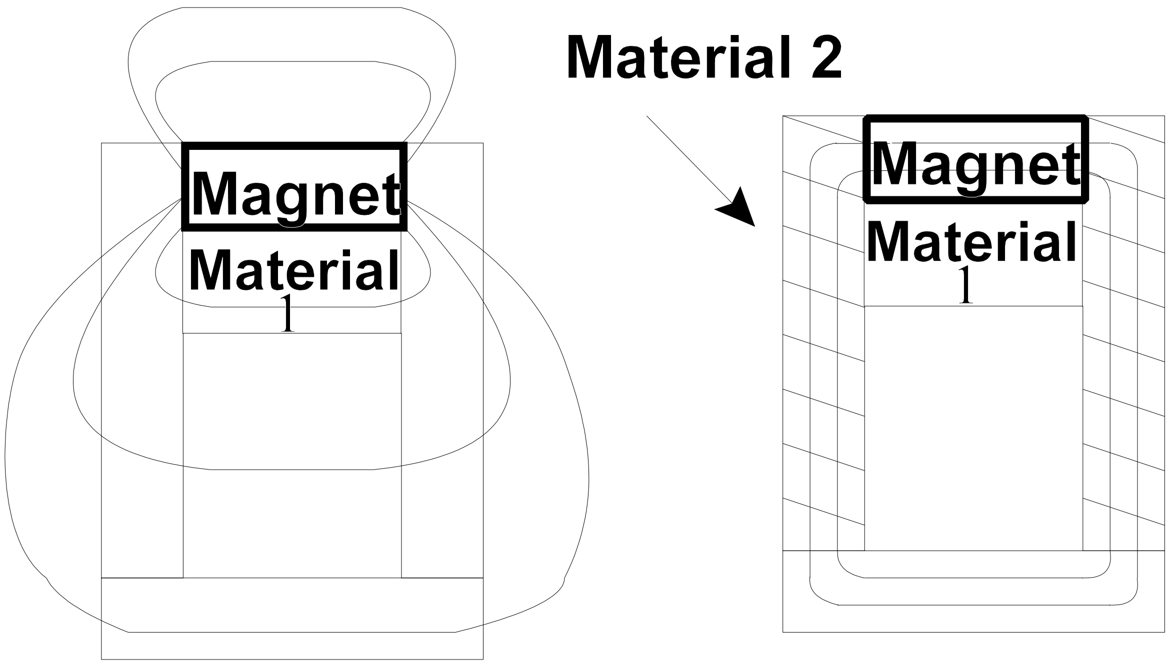


Figure 14 Lines of Force in Magnet Holder Experiment

Note in Fig. 14, that the number of lines available on the left is very small due to the distance between the magnet and the target, and because of the orientation of the magnet - we get more lines of force at the ends of the magnet than at the sides. The number of lines is concentrated, however, by the presence of material 2 in the holder on the right, as material 2 bends the lines of force. This results, of course, in a much greater force. Materials like material 2 are called ferro-magnetic materials and include iron, nickel, cobalt, and various alloys like alnico (aluminum, nickel, cobalt), and various rare-earth alloys like neodymium.

Now, so far our experiments have shown nothing that would prevent us from simply using the magnetic field intensity H as our characterization of the magnetic force. There is, however one last experimental result that requires us to abandon H and introduce a new concept to describe magnetic forces. If we wind two identical electromagnets as shown in Fig. 15: they have the same number of turns, they have the same length, they have the same current flowing in their windings; we find that they produce vastly different forces AT THE SAME DISTANCE. More force means more "lines of force," but the same current, number of turns, and length means that both magnets have the same magnetic field intensity H.

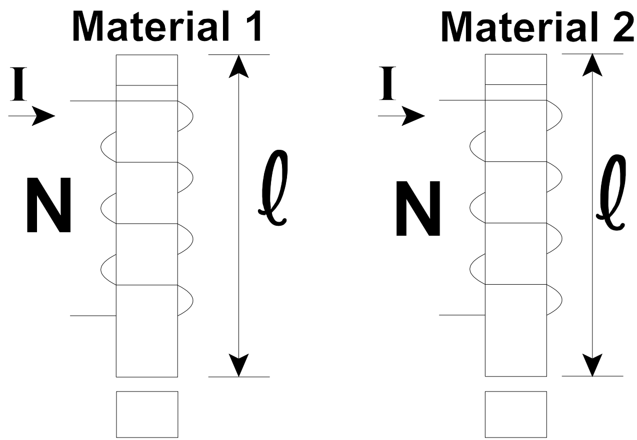
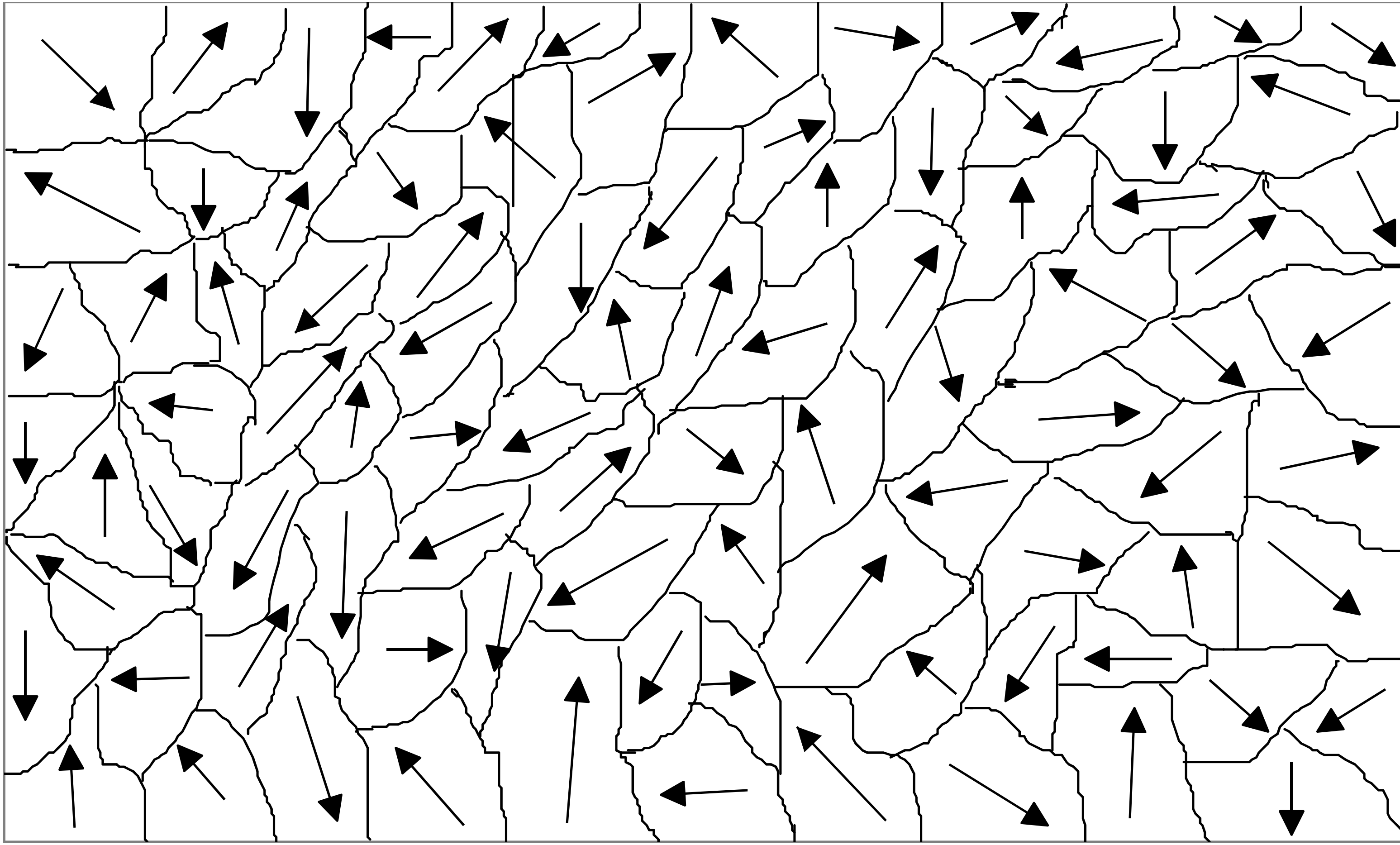


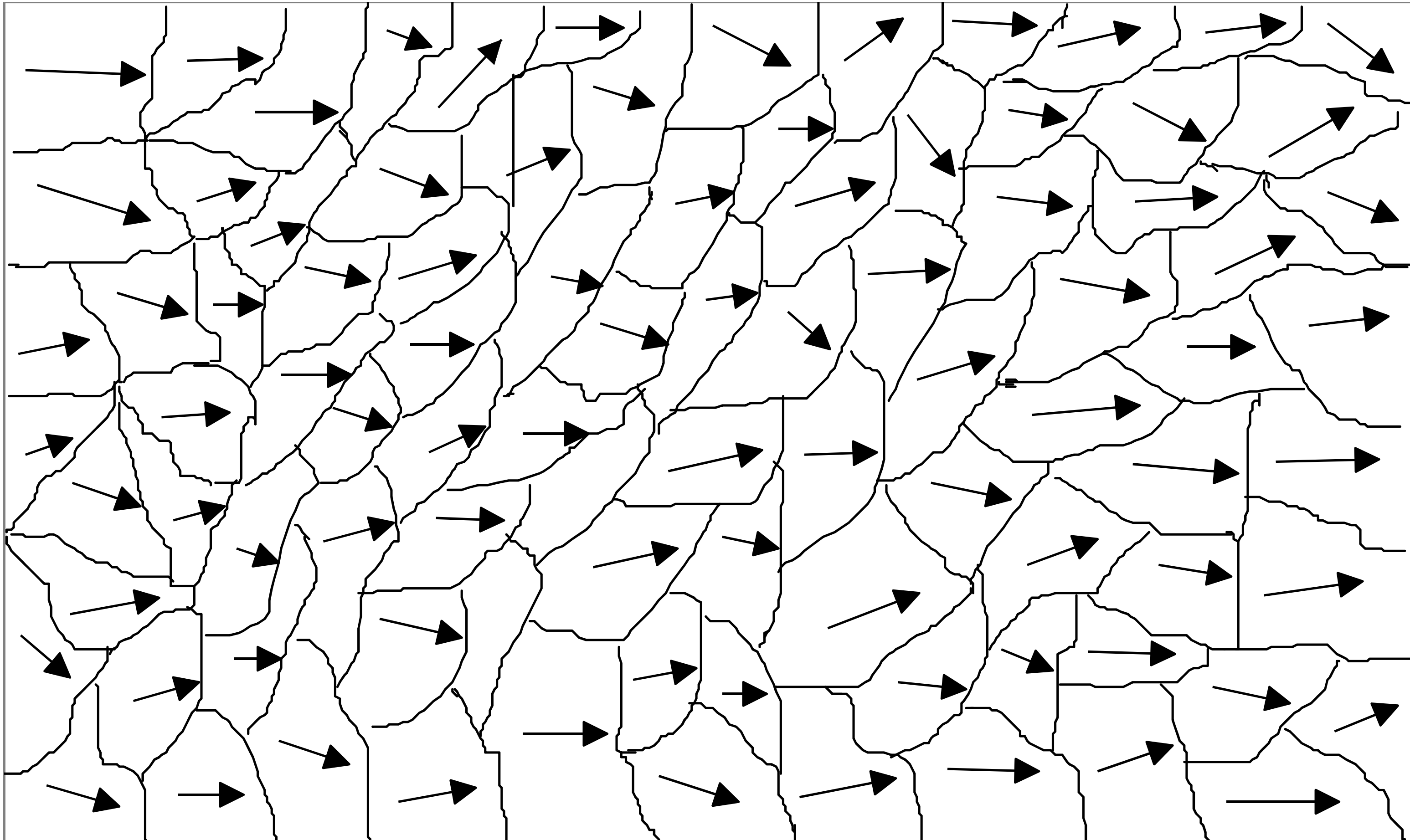
Figure 15 Electro-magnets Wound on Different Core Materials

Ferro-magnetic materials are poly-crystaline substances whose individual crystal structures contain magnetic domains. A magnetic domain is a group of atoms that contain spin-coupled electrons (notably unpaired electrons in the 3d shell). These spin-coupled electrons give the magnetic domain a definite magnetic orientation as shown below. (For convenience in drawing we have shown each separate crystal structure as containing only one magnetic domain. In reality each separate crystal structure contains many randomly oriented magnetic domains.)



Iron unmagnetized - domains magnetized in random directions

When an external magnetic field is applied to a ferro-magnetic material, its domains tend to line up with the external field as shown below.



Iron strongly magnetized - domains magnetized in roughly the same direction

The aligned domains produce lines of force in addition to the flux density $B = \mu_0 \cdot H$

When an electromagnet is wrapped around a ferro-magnetic core, the flux density produced is given as

$B = \mu_0 \cdot H + M$ where M is the "extra" flux induced in the core by the magnetic field.

To resolve this difficulty, we define B, the magnetic flux density as the number of lines per unit area, and relate B and H as

$$B = \mu \cdot H$$

where μ is a parameter of the material called the permeability. The permeability is a measure of how easily lines of force (flux lines) can travel through a material. The permeability of non-ferromagnetic materials is approximately the permeability of free space which is called μ_0 and has the value

$$\mu_0 = 4 \cdot \pi \cdot 10^{-7}$$

in SI (MKS) units. Very often the permeability μ is described by the relative permeability μ_R where

$$\mu = \mu_R \cdot \mu_0$$

Units

	SI	cgs	U.S. Customary
Magnetic Flux ϕ	Webber	Maxwell	Lines (or kilo-lines)
Magnetic Field Intensity H	Amp Turns per meter	Oerstead	Amp Turns per inch
Magnetic Flux Density B	Tesla	Gauss	Lines per square inch (or kilo-lines per square inch)

The conversion factors are

$$1\,\text{Webber} = 10^8\,\text{Maxwell} = 10^5\,\text{kilo-line}$$

$$1\,\frac{\text{Amp Turn}}{\text{m}} = 1.257\cdot 10^{-2}\,\text{Oerstead} = 2.540\cdot 10^{-2}\,\frac{\text{Amp Turn}}{\text{in}}$$

$$1\,\text{Tesla} = 10^4\,\text{Gauss} = 64.52\,\frac{\text{kilo-line}}{\text{in}^2}$$

In The Lab

Do an internet search for Beakman's motor and follow his instructions to construct the motor.

(The next exercise is best done as a team- project outside of class.)

Wrap 200 feet (approximately 3000 turns) of #30 magnet wire on the iron core and on the brass core, try to get the magnets as much alike as possible. For each magnet connect the 9V battery, measure the current and the maximum weight that the magnet can lift.
Since the force (weight that can be lifted) is proportional to B, we have

$$\frac{F_{Steel}}{F_{Brass}} = \frac{B_{Steel}}{B_{Brass}}$$

Since B and H are related as

$$B = \mu \cdot H$$

we have

$$\frac{F_{Steel}}{F_{Brass}} = \frac{B_{Steel}}{B_{Brass}} = \frac{\mu_{Steel} \cdot H_{Steel}}{\mu_{Brass} \cdot H_{Brass}}$$

Since H is given as

$$H = \frac{N \cdot I}{\ell}$$

we have

$$\frac{F_{Steel}}{F_{Brass}} = \frac{B_{Steel}}{B_{Brass}} = \frac{\mu_{Steel} \cdot \frac{N \cdot I_{Steel}}{\ell}}{\mu_{Brass} \cdot \frac{N \cdot I_{Brass}}{\ell}} = \frac{\mu_{Steel} \cdot I_{Steel}}{\mu_{Brass} \cdot I_{Brass}}$$

so we may calculate the relative permeability of the iron as (the permeability of brass is assumed to be μ_0)

$$\mu_{Steel} = \frac{F_{Steel}}{F_{Brass}} \cdot \frac{I_{Brass}}{I_{Steel}} \cdot \mu_{Brass}$$

Note that even though the coils are as identical as we can make them, there are several factors that prevent us from achieving the same current in both coils. The most important effect is that both the open circuit voltage and the internal resistance of a battery change as the battery discharges. As the battery becomes weaker, the open circuit voltage goes down, and the internal resistance increases. This, coupled with the fact that the resistance of the coil depends on the tension in the wire as it is being wound, means that it is not possible for us to get the same current in each coil. Luckily it is easy to take this into account as shown above.
Note that the resistance of a wire is given as

$$R_{Wire} = \frac{\rho \cdot \ell}{A}$$

where ρ is a constant of the material of which the wire is made called the resistivity, ℓ is the length of the wire, and A is the cross-sectional area of the wire. As you increase the tension in the wire as you wind, the wire gets longer and the cross-sectional area gets smaller.

Hints:
Before you begin winding, file the ends of both cores as smooth and perpendicular as you can make them.

Use a variable-speed electric drill to wind the magnets. You must be extremely careful to lay the windings down in very even layers - your success will depend a very great deal on the level of your craftsmanship.

Get the winding as close to one end of the magnet as possible. (This is the end you will use to test the force.)

Do NOT measure the force by sticking the magnet into a box of nails, paperclips, etc. The force further from the magnet will be weaker. Instead find a thin, square metal plate; lift the plate with the magnet; and load the nails or paperclips on the plate.

If you are unable to measure the force of the brass-core magnet, use 50 mg as the mass that the brass magnet can lift. (This is the mass my brass magnet could lift.)

Motors and Generators

Consider a current-carrying wire in a magnetic field as shown in Fig.1.
The Magnetic field points down, and the force is toward the back (into the page).

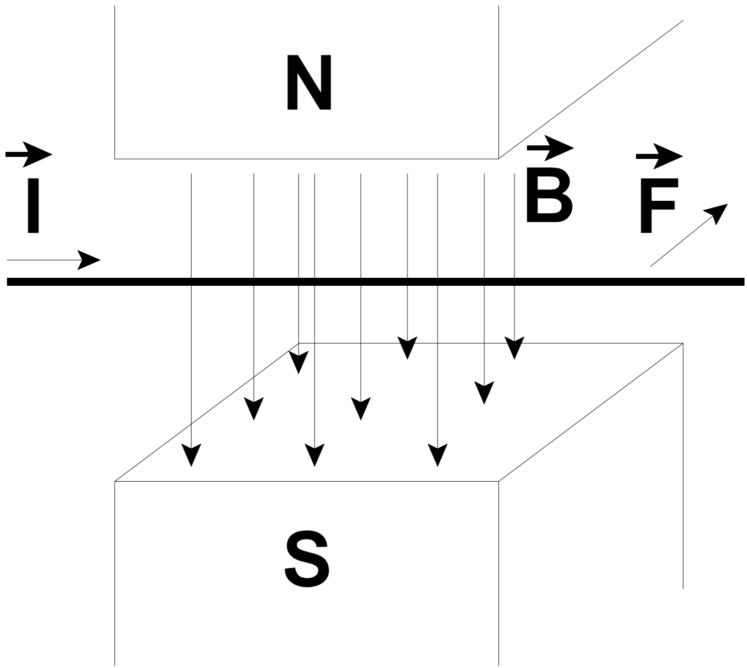


Figure 1 Current-carrying Wire in a Magnetic Field B

The wire experiences a force away from the reader, i.e., into the page. This force can be described by the vector equation

$$\overrightarrow{F} = q\vec{v} \times \overrightarrow{B}$$

where $\overrightarrow{F}$ is the force vector, $\vec{v}$ is the velocity vector of the moving (positive) charges that constitute the current $\vec{I}$, $\overrightarrow{B}$ is the magnetic field vector (the magnetic flux density as previously defined, and $\times$ is the vector cross product.

The force can also be visualized by an examination of the magnetic force lines of the wire and the imposed field B as shown in Fig. 2 on the next page. Fig. 2 is a right-hand view of Fig.1, i.e., Fig.2 is a two-dimensional view of Fig. 1 as viewed from the right. So the front of Fig. 1 is at the left in Fig. 2, and the back of Fig.1 is at the right in Fig. 2.

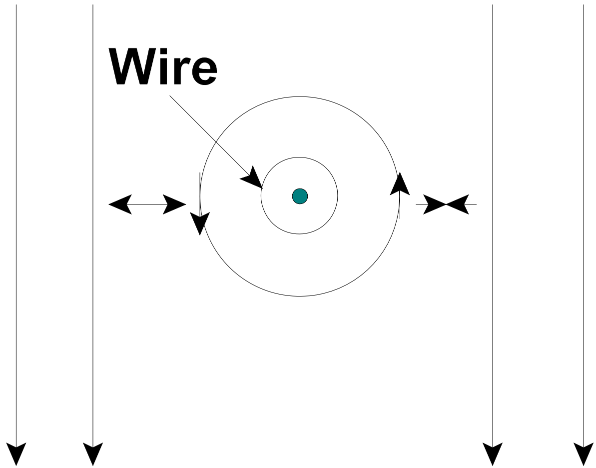


Figure 2 Current-carrying Wire in a Magnetic Field B
(Right Side View)

This is essentially the mechanism that makes the Beakman motor operate. To see the details of the operation of such a motor we study Fig. 3.

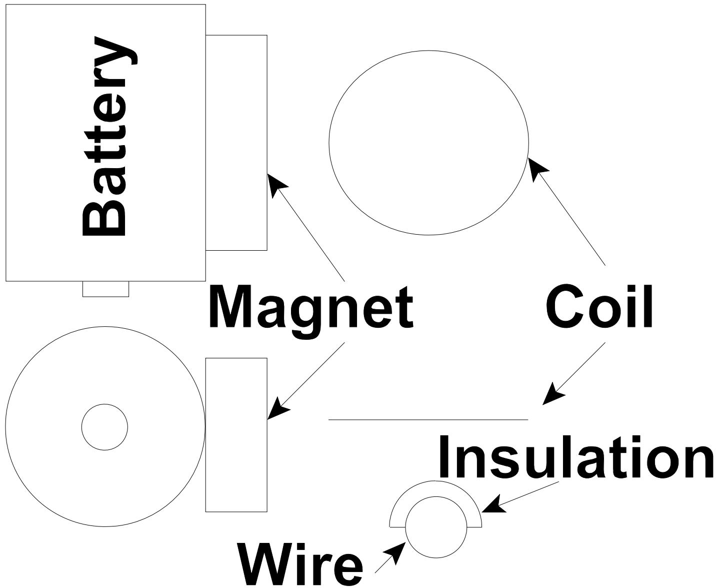


Figure 3 Beakman's Motor: Parts Identification

Note in Fig. 3, that the half-insulated wire that serves as the commutator rotates as the coil rotates. The operation of the motor is shown in Figs. 4 - 9. Fig. 4 shows the general principle of operation where the motor is producing maximum clock-wise torque as the north pole above the coil is repelled by the north pole of the permanent magnet, and the south pole below the coil is attracted to the north pole of the permanent magnet.

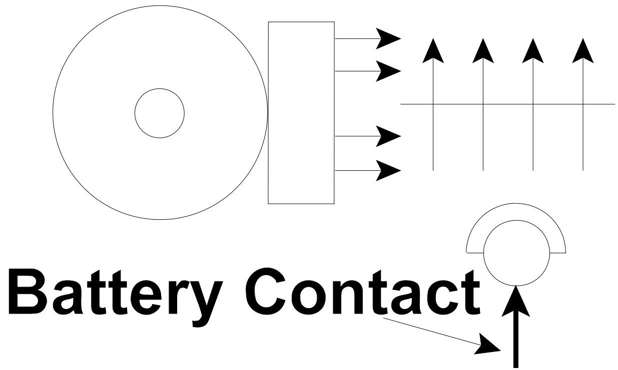


Figure 4 Maximum Clock-wise Torque

Figs. 5 - 7 show the rotor in several successive positions. Fig. 5 shows the motor in its initial position before being tapped into operation. Notice that the insulation prevents contact between the coil and the battery so there is no current in the coil, and, therefore, no magnetic field produced by the coil.

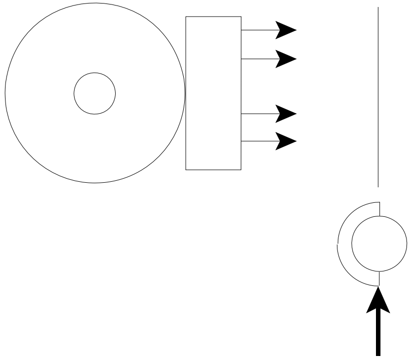


Figure 5 Initial Position: Before Rotation

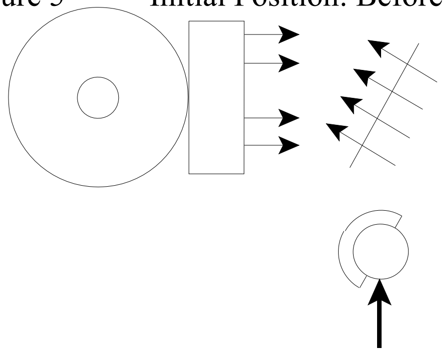


Figure 6 Start of Rotation

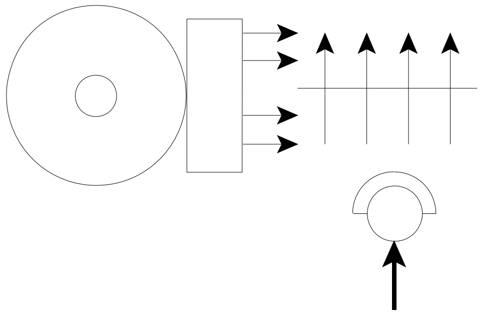


Figure 7 Maximum Clockwise Torque

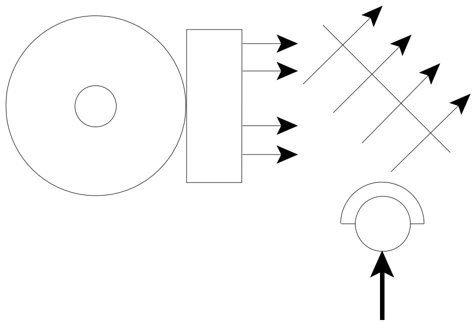


Figure 8 Near the End of Torque-producing Portion of Rotation

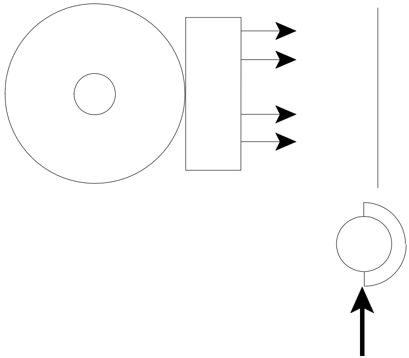


Figure 9 End of Torque-producing Portion of Rotation

Notice that in Fig. 9 there is no current in the coil and, therefore, to torque being produced. One might be tempted to ask ,“Why does the motor continue to rotate with no torque applied to the coil?” To answer that question we must review the notion of inertia. The first two of Newton’s laws of motion describe and quantify inertia. The first law states: “A body in rectilinear (straight-line) motion tends to stay in motion unless acted on by an outside force. A body at rest tends to stay at rest.” The second law states that the force required to accelerate a body by an amount a (with units $\frac{m}{s^2}$) is

$$\overrightarrow{F} = m \cdot \overrightarrow{a}$$

where the vector notation reminds us that the force must be in the direction of the desired acceleration, and that an acceleration is necessary to effect a change of direction. (Velocity has both magnitude and direction.) It is the mass that is responsible for inertia. The notion of inertia can be generalized to the case of rotating bodies as

$$\vec{T} = J \cdot \vec{\alpha}$$

where T is the torque, α is the rotational acceleration(with units $\frac{radians}{s^2}$),

and J is the rotational inertia. Rotational inertia is the phenomenon that keeps the coil spinning for several seconds if you build the motor without removing the insulation from the coil wire and attempt to tap the coil into rotation. It is the rotational inertia of the coil that keeps it rotating during the time that the insulation covered portion of the coil wire is in touching the battery contact (paper clip).

One might ask: “why leave any insulation at all?” To answer this question consider Fig.10.

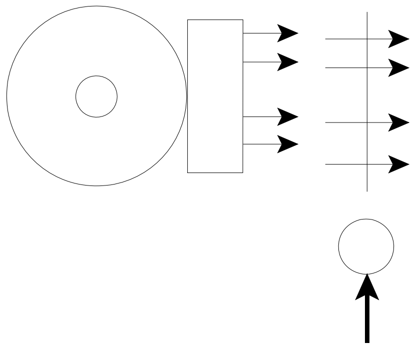


Figure 10 Motor with Insulation Removed All the Way Around

Notice that in Fig. 10, the north pole of the permanent magnet uniformly attracts the south pole of the coil and the coil would stop (and lock) in this position. The partially stripped coil wire acts as a commutator allowing current to flow only when and only in the direction we desire.

There are two very big problems with the Beakman motor: these problems are so serious that they render the Beakman motor no more than a (reasonably instructive) toy. The first problem is that current flows during only one half of the rotation, and the other problem is that the magnetic flux lines flow through low-permeability air instead of high-permeability iron or steel.

The first problem is solved with the split-ring, graphite-brush commutator.

Consider the machine of Fig. 11

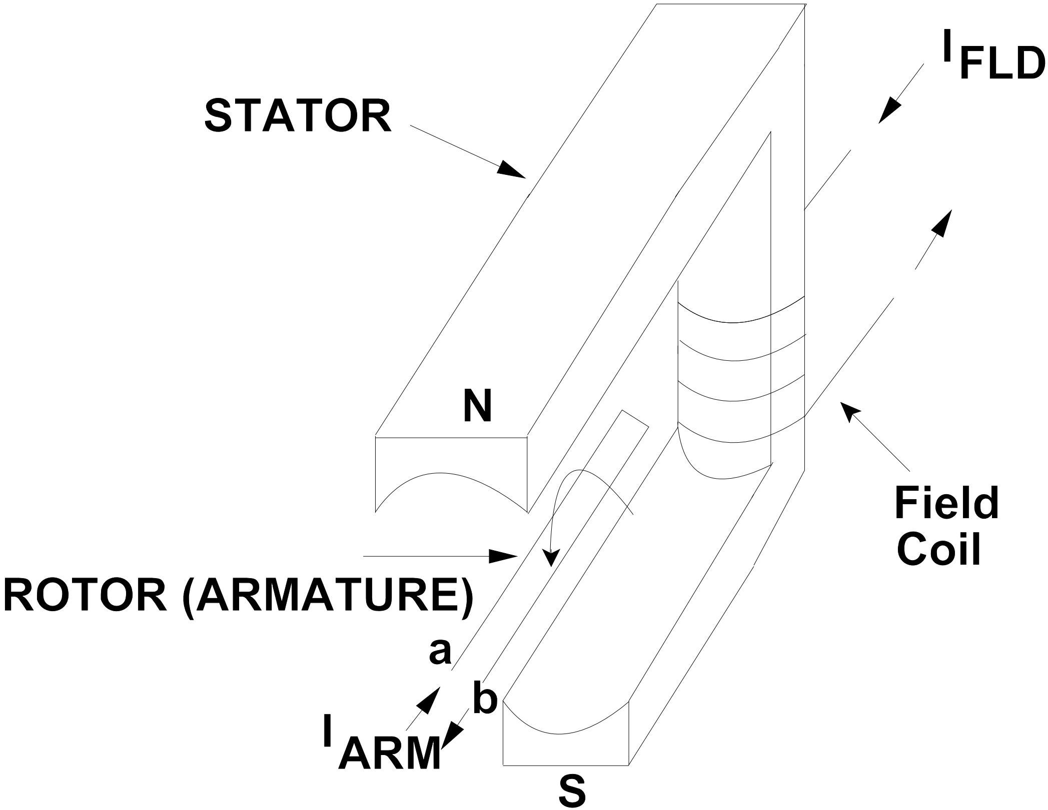


Figure 11 Simple DC Machine operated as a motor

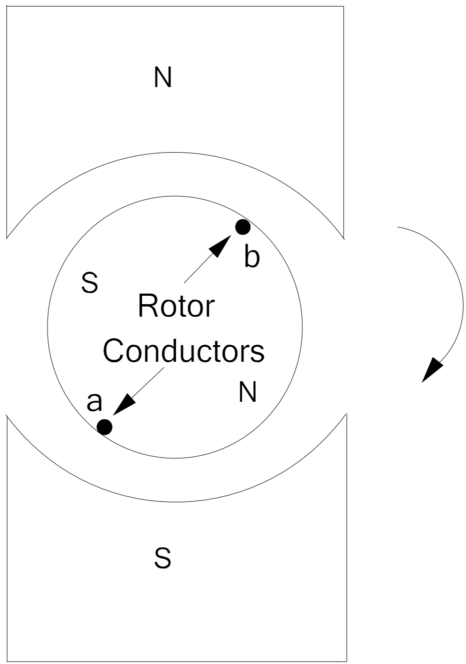


Figure 12   Front view of Figure 11

If we excite the machine of Fig. 11 as shown, the armature would rotate CCW until a and b were in the horizontal plane, and stop as shown in Fig. 13.

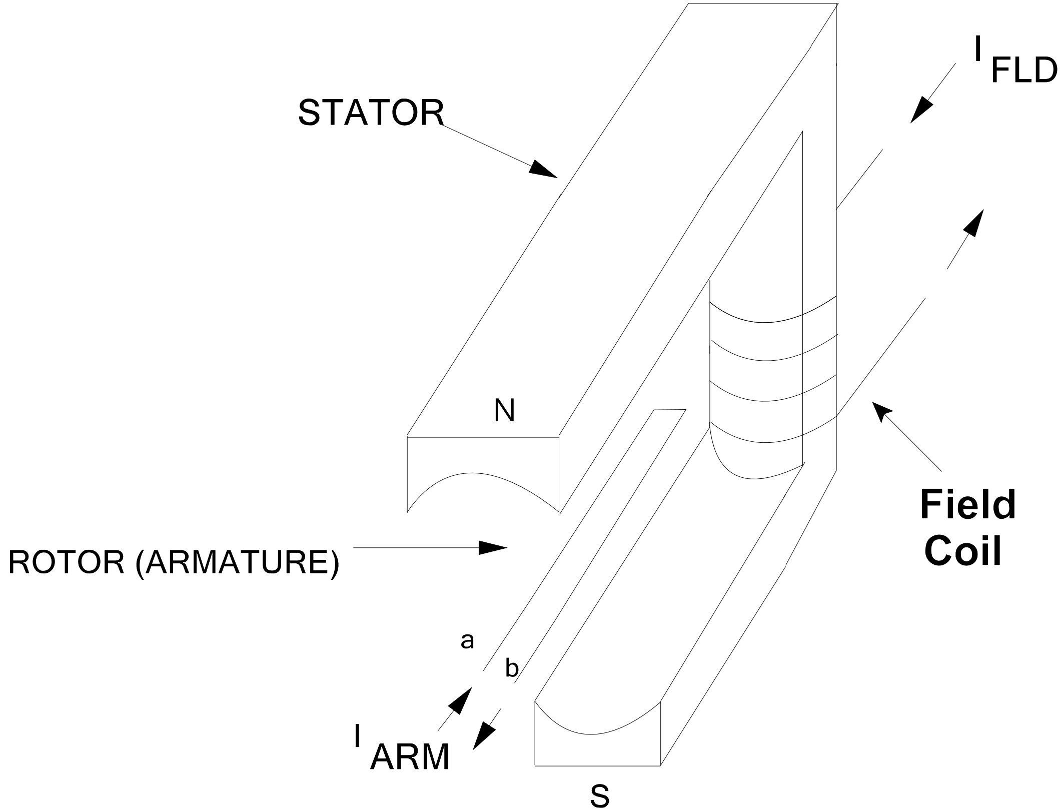


Figure 13   Simple DC Machine with Rotor Frozen in the Horizontal Plane

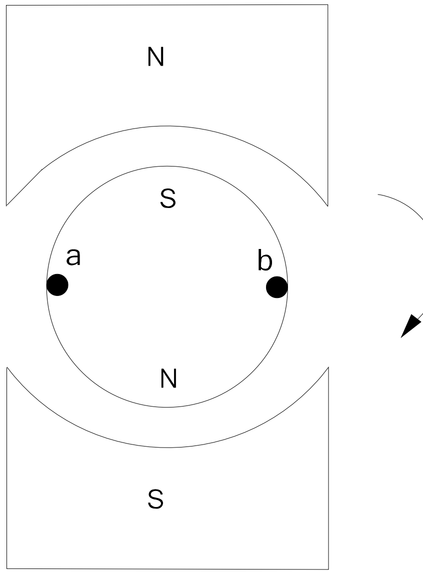


Figure 14 Front View of Figure 13

This happens because the north pole of the armature is attracted to the south pole of the field coil, and the south pole of the armature is attracted to the north pole of the field coil. But, suppose that at this point we could reverse the flow of armature current so that what was the armature north pole becomes the armature south pole and vice versa. Then the armature would continue to rotate as the north pole of the armature is repelled by the north pole of the field coil, and the south pole of the armature is repelled by the south pole of the field coil.

Consider the modification of the machine of Fig. 11 shown in Fig. 15.

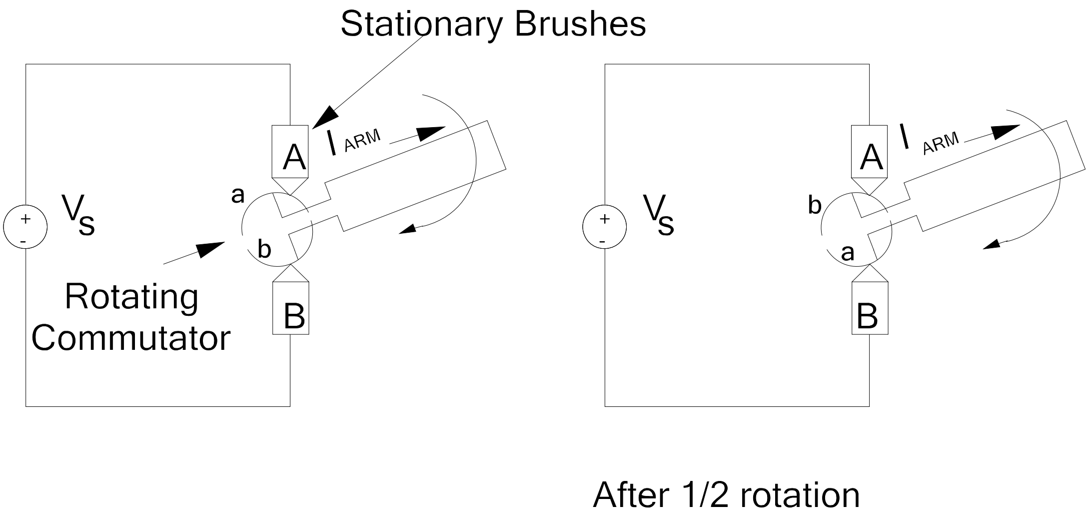


Figure 15 Split-ring, Graphite-brush Commutator

As the commutator rotates, the voltage at point ab changes direction twice per revolution, reversing the rotor current (and therefore the direction of the rotor magnetic field), thereby causing the motor to rotate continuously in the counter-clockwise direction. (Remember that this is a modification to the machine of Fig. 11). The magnitude of the voltage at point ab determines the current in the rotor which determines the strength of the rotor magnetic field, which determines the torque developed by the motor. In an unloaded motor a change in the torque developed by the motor results in a change in speed. Within the limits of the motor specification, increasing the rotor voltage increases the speed and decreasing the rotor voltage decreases the speed. If the polarity of the voltage Vs is reversed, the motor rotates in the opposite direction. Direction reversal may also be accomplished by reversing the current in the stator (field) winding. Increasing the field current DECREASES the speed and decreasing the field current INCREASES the speed. This is why, when separately excited D.C. machines are de-energized, the armature current is always returned to zero first, then the field current is reduced to zero. If the field current were reduced first the rotor would "run away", possibly damaging the motor (and anyone standing nearby). To continue our study of the dc machine we must understand Faraday’s Law: Faraday’s Law explains the operation of electrical transformers, generators and motors. Consider the situation from Fig. 1, but with no current injected in the armature and the wire possessing a velocity toward the back of the page as shown in Fig. 16.

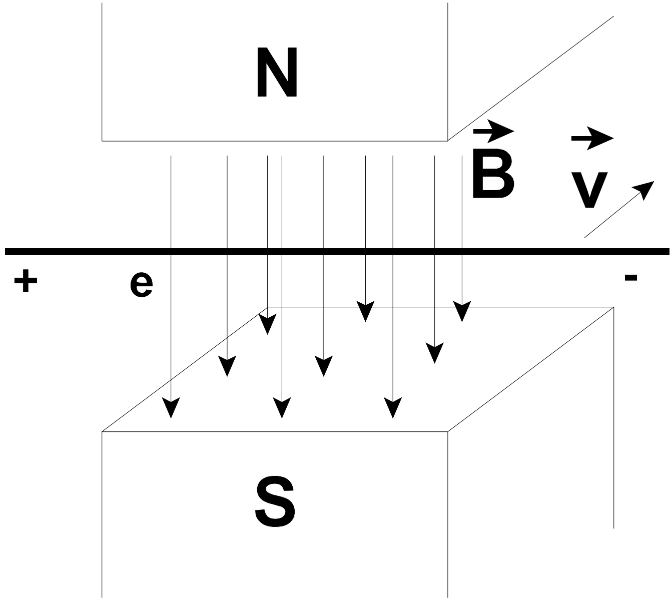


Figure 16   Faraday's Law

Faraday's Law says that a conductor that experiences a changing magnetic field develops a voltage (electro-motive force) e across the conductor. Notice that the magnetic field could be changing because the conductor is moving through a stationary magnetic field (as in Fig. 14), or because a magnetic field is moving with respect to a stationary conductor.

Mathematically Faraday’s Law is stated as:

$$e = \frac{\partial \vec{B}}{\partial t}$$

So, if one disconnects the armature supply from a D.C. machine, leaves the field current intact, and mechanically spins the rotor, a voltage is developed across the armature terminals in accordance with Faraday's law. This operation of the machine is referred to as running the machine as a generator. As the generator terminals are electrically loaded (i.e. a load resistance is connected across the generator terminals, (causing a load current to flow), the machine presents a higher torque resistance at its shaft. (The shaft becomes harder to rotate.) Of course, in all these examples the field coil could be replaced with a permanent magnet with no loss of generality.

This brings up an interesting point; If the wire in Fig. 1 moves because of the current flowing in it, then as the wire moves through the magnetic field a voltage (called the back emf, or the speed voltage) e is induced in the wire. This fact coupled with the fact that the wire used to wind the armature coil has resistance

$R = \frac{\rho \ell}{A}$ as we have seen in the chapter on magnetics,

leads us to the circuit model of the DC machine as shown in Fig 17.

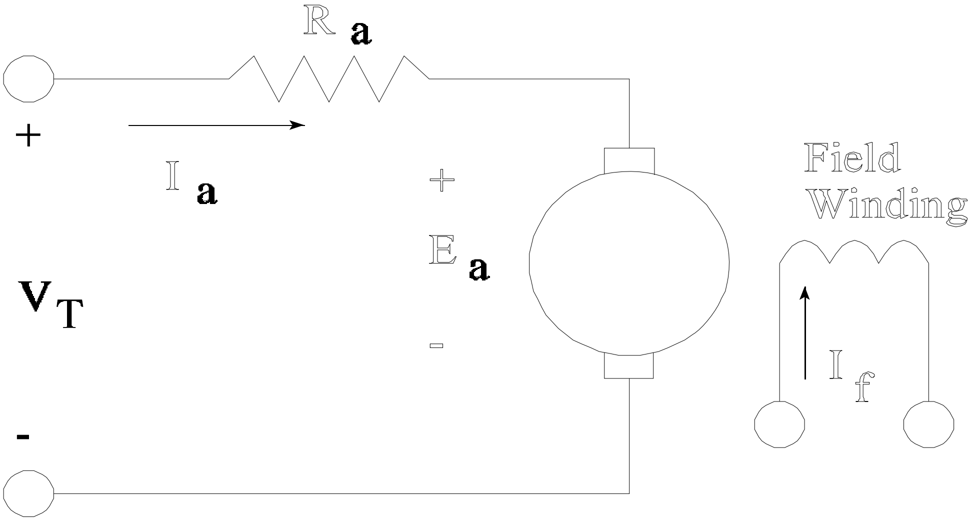


Figure 17 DC Machine SteadystateCircuit Model

As we have seen, if we produce a magnetic field in the stator of a D.C. machine by causing curent in the field winding, we can cause the machine to behave either as a motor or as a generator depending on what we do with the armature circuit. If we apply current to the armature the rotor begins to spin (like poles repel - unlike poles attract) and we have a D.C. motor. If , on the other hand, we drive the rotor mechanically the machine produces electricity in its armature circuit (by Faraday induction) and we have a generator.

It is important to notice that even when we operate the machine as a motor, as the rotor spins a voltage is induced in the armature coil (which we are driving with a voltage source to cause motor action), and that this induced voltage opposes the voltage we are trying to use to spin the rotor. In fact, it is this "back emf" not the armature resistance that primarily determines (along with the magnitude of the applied voltage) the armature current in a D.C. motor. The D.C. machine (whether it is operated as a motor or a generator) is thought of as having the circuit model of Fig. 15.

The model has two electrical ports, the armature port and the field port, each of which requires two terminals. R_a represents the D.C. resistance of the armature coil, V_T is the armature terminal voltage, I_a is the armature current, I_f is the field current, R_f is the D.C. resistance of the field winding, and E_a is the back emf which is given in terms of the machine constants as:

$$E_a = K_a \phi \omega$$

where K_a is a constant that depends on the stator core material and the construction of the machine, ϕ is the core flux, and ω is the rotational speed of the rotor in radians per second.

The speed ω in radians per second can be related to the speed n in revolutions per minute (RPM) as:

$$\omega = \frac{\pi}{30} \cdot n$$

The armature current I_a is given then as:

$$I_a = \frac{V_T - E_a}{R_a}$$

Since the flux ϕ depends on the field current I_f,the D.C. machine can be conveniently characterized by a plot of E_a vs. I_f at constant (usually rated) speed. Such a plot is referred to as the magnetization curve of the machine and a sample is shown below in Figure 18.

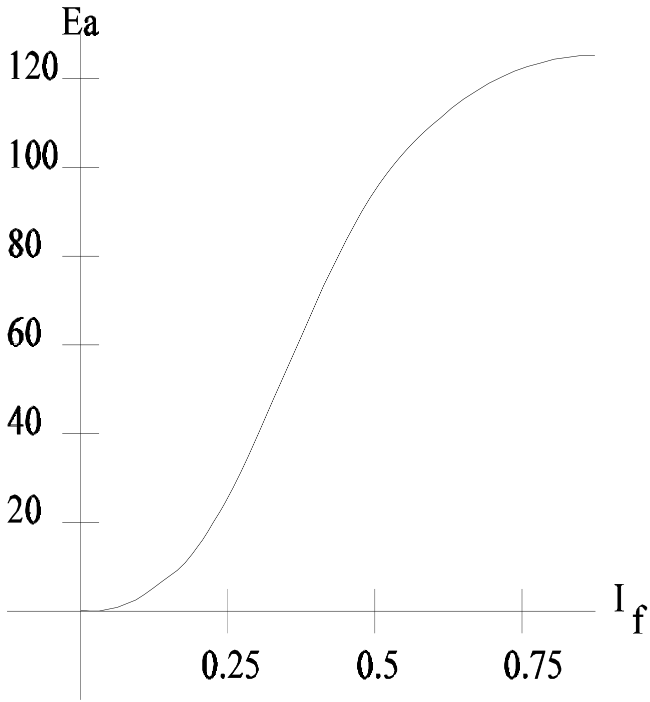


Figure 18   DC Machine Magnetization Curve

Notice the similarity of this curve to the B H curve of ferro-magnetic materials. This similarity is not accidental, since the magnetic characteristics of the core material are what cause the magnetization curve of the D.C. machine. The magnetic field strength H is directly proportional to I_f (the field current), and the terminal voltage Ea is proportional to ϕ which is proprotional to B (the magnetic flux density). The Ea vs. I_f curve then, has the same shape as the B vs. H curve.

You will notice that there is space for two machines with their shafts coupled together at each station. The machine on the left is the "Machine Under Test" i.e., the machine whose characteristics we are interested in studying. The machine on the right is a D.C. machine called the dynamometer, whose purpose is to excite the machine under test either electrically or mechanically for the purpose of carrying out the test. As we have seen the D.C. machine (indeed any of the electric machines we shall study) can be operated as either a generator or a motor. When the experiment requires that the rotor of machine under test be rotated, the dynamometer is operated as a motor whose rotating shaft rotates the shaft of the machine under test. When the experiment requires that a mechanical torque load be applied to the shaft of the machine under test, the dynamometer is operated as a generator loaded by the power resistor bank. The dynamometer shaft then, resists being twisted and thus presents a mechanical torque load to the machine under test. The coupling of the shafts is usually indicated on the schematic by a dashed line connecting the two armatures together.

1) With all breakers open and all D.C. power supply voltage control knobs in the fully counter clockwise position, construct the circuit shown in Figure 19.

2) Close the shunt field supply breakers, and slowly raise the shunt field current to 0.3 Amperes.

3) Close the armature supply breakers and the D.C. machine breaker, and slowly raise the armature voltage to 25 Volts.

4) Make sure the tachometer sensor is pointed properly at the right end of the D.C. machine shaft and record the shunt field voltage, the armature current, and the speed, noting the direction of rotation.

5) Reduce the armature supply to zero, then reduce the shunt field supply to zero, and open all breakers.

6) Reverse the connections to the shunt field winding.

7) Close the breakers, bring the shunt field current to 0.3 A, then bring the armature voltage to 25 V, and record the shunt field voltage, the armature current and the speed, noting the direction of rotation.

8) Reduce the supplies to zero, (armature first, then shunt field), open the breakers, return the shunt field windings to their original configuration, reverse the leads to the armature winding, and repeat the test procedure.

9) Restore the original configuration, bring the shunt field to 0.4 A, and set the armature voltage 50 V and record the armature current and the speed. Increase the armature voltage to 100 V in 10 volt increments, recording armature current and speed at each step.

10) Repeat 9) for shunt field supply = 0.35 A, 0.3 A and 0.25 A.

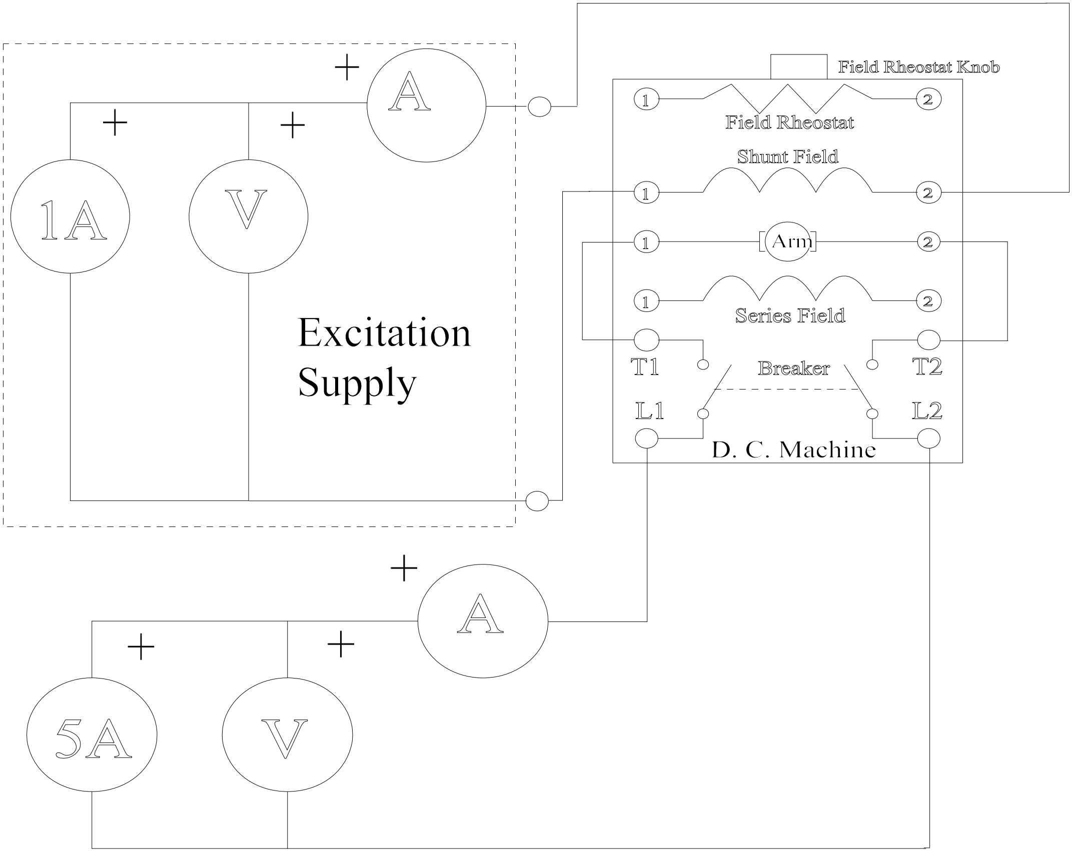


Figure 19 Motors and Generators Lab

Logic and Binary Numbers

Boolean algebra is the name given to the study of the truth or falsity of given propositions. The variables in Boolean algebra can take on only two discrete values: true and false. These are often notated as zero (or low) for false and one (or high) for true. The allowable values for the low voltage and the high voltage vary from logic family to logic family. Two of the most common logic families are TTL (Transistor-Transistor-Logic) and CMOS (Complimentary-Metal-Oxide-Silicon). For both these families the low voltage is nominally zero; the high voltage for TTL is nominally 5 V; and the high voltage for CMOS is nominally any where between 3V and 18V, depending on the power supply available. The same logic functions are available for both families, but the pin-outs are different. Several functions are defined for Boolean algebra as shown below. A Boolean function is defined by its truth-table, which lists all possible combinations of inputs and whether the function is true or false for each. For example:

And

Inputs		Output
x	y	z
0	0	0
0	1	0
1	0	0
1	1	1

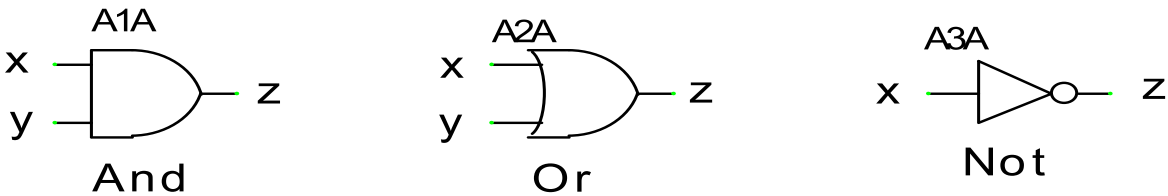
Or

Inputs		Output
x	y	z
0	0	0
0	1	1
1	0	1
1	1	1

Not (Inverter)

Input	Output
x	z
0	1
1	0

These functions are symbolized by logic-gates as shown below.



The Nand function is the same as an And followed by a Not, and a Nor is the same as an Or followed by a Not.

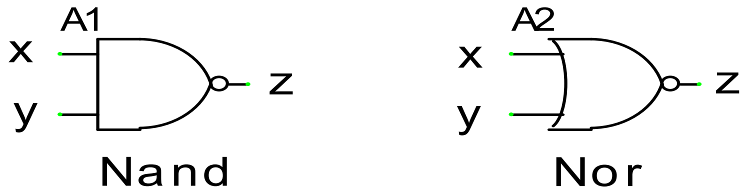
Nand

Inputs		Output
x	y	z
0	0	1
0	1	1
1	0	1
1	1	0

Nor

Inputs		Output
x	y	z
0	0	1
0	1	0
1	0	0
1	1	0

These functions are symbolized by logic-gates as shown below.



Note that in all cases the inversion is carried by the circle at the output.

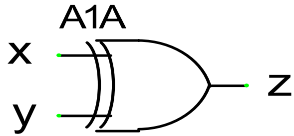
Note that the Or function is true if either input or both inputs are true; i.e. the true state includes the situation where both inputs are true. There is another flavor of Or gate called an Exclusive Or (Xor) that is true if either input is true, but not if both inputs are true; i.e. the true state excludes the situation where both inputs are true. An Exclusive Or followed by a Not is called an Exclusive Nor gate (Xnor). The truth-tables and gate representations are shown below.

Xor

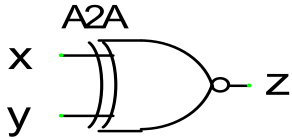
Inputs		Output
x	y	z
0	0	0
0	1	1
1	0	1
1	1	0

Xnor

Inputs		Output
x	y	z
0	0	1
0	1	0
1	0	0
1	1	1



Xor



Xnor

Binary Numbers

The order in which the inputs are listed in truth-tables is not arbitrary: to understand the ordering we need to study a separate but closely related problem. The problem is how to represent numbers with not the usual ten digits, but with two digits: 1 and 0. Such numbers are called binary numbers.

To better understand how this might be accomplished, let us review what we know about our usual decimal number system.

0	10	20	30	40	50	60	70	80	90	100
1	11	21	31	41	51	61	71	81	91	101
2	12	22	32	42	52	62	72	82	92	102
3	13	23	33	43	53	63	73	83	93	103
4	14	24	34	44	54	64	74	84	94	104
5	15	25	35	45	55	65	75	85	95	105
6	16	26	36	46	56	66	76	86	96	106
7	17	27	37	47	57	67	77	87	97	107
8	18	28	38	48	58	68	78	88	98	108
9	19	29	39	49	59	69	79	89	99	109

Note that when we run out of digits, we place a zero in the column and start a new column to the left. Decimal numbers are represented in what is called a positional notation. The meaning of the number is conveyed not only by the values of the digits but by their position in the number. For example, 1524 means 4 x 1 + 2 x 10 + 5 x 100 + 1 x 1000, whereas
4215 means 5 x 1 + 1 x 10 + 2 x 100 + 4 x 1000. We note that $1000 = 10^3$, $100 = 10^2$, $10 = 10^1$, and we define $x^0 = 1$, so $1 = 10^0$. Then we have,

$$1524 = 4 \cdot 10^0 + 2 \cdot 10^1 + 5 \cdot 10^2 + 1 \cdot 10^3$$

and

$$4215 = 5 \cdot 10^0 + 1 \cdot 10^1 + 2 \cdot 10^2 + 4 \cdot 10^3$$

The total number of different digits is called the base of the number system and for a different number of digits, we would replace 10 with the appropriate number of digits. Consider counting with only two digits we write

0
1

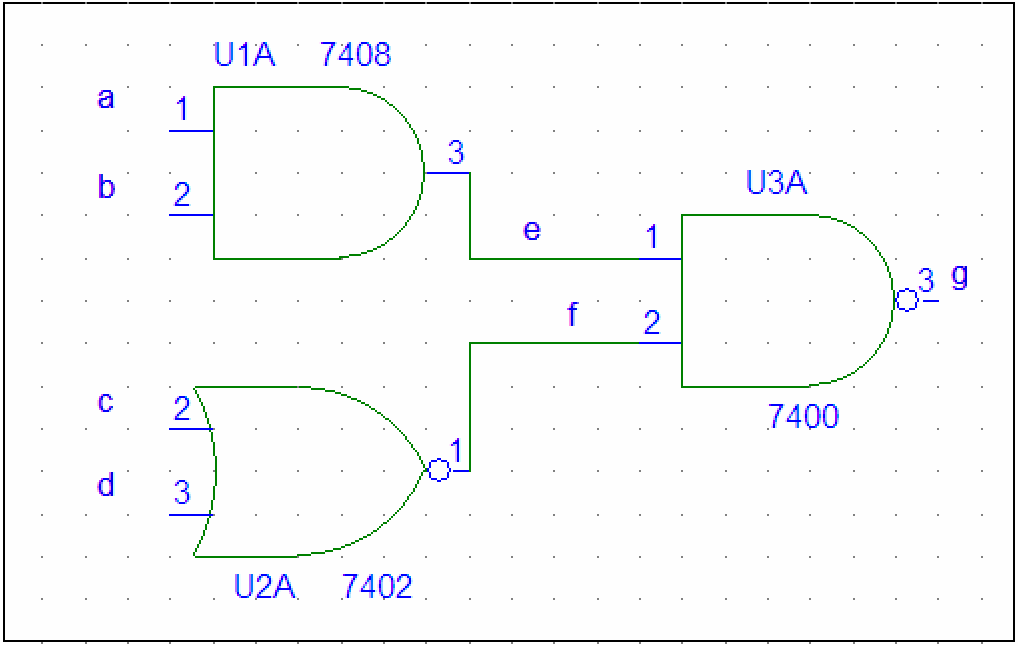
and we have already run out of digits, so we place a zero in the column and start a new column as shown below

```
 0
 1
00
01
10
11
```

We have run out of digits again so we start a new column.

Binary	Decimal		Binary	Decimal		Binary	Decimal		Binary	Decimal
0	0		10000	16		100000	32		110000	48
1	1		10001	17		100001	33		110001	49
10	2		10010	18		100010	34		110010	50
11	3		10011	19		100011	35		110011	51
100	4		10100	20		100100	36		110100	52
101	5		10101	21		100101	37		110101	53
110	6		10110	22		100110	38		110110	54
111	7		10111	23		100111	39		110111	55
1000	8		11000	24		101000	40		111000	56
1001	9		11001	25		101001	41		111001	57
1010	10		11010	26		101010	42		111010	58
1011	11		11011	27		101011	43		111011	59
1100	12		11100	28		101100	44		111100	60
1101	13		11101	29		101101	45		111101	61
1110	14		11110	30		101110	46		111110	62
1111	15		11111	31		101111	47		111111	63

We note that the logic inputs are listed in binary order in our truth tables.
The gates may be combined, and the overall truth table determined as follows:



We build the overall truth table as follows. First we note that there are four inputs: a, b, c, and d, so we write all possible combinations of the four inputs as follows.

a	b	c	d	e	f	g
0	0	0	0			
0	0	0	1			
0	0	1	0			
0	0	1	1			
0	1	0	0			
0	1	0	1			
0	1	1	0			
0	1	1	1			
1	0	0	0			
1	0	0	1			
1	0	1	0			
1	0	1	1			
1	1	0	0			
1	1	0	1			
1	1	1	0			
1	1	1	1			

Next we use the truth tables of the “And” and “Nand” gates to fill in columns e and f as shown below.

e is a And b, f is c Nor d.

a	b	c	d	e	f	g
0	0	0	0	0	1	
0	0	0	1	0	0	
0	0	1	0	0	0	
0	0	1	1	0	0	
0	1	0	0	0	1	
0	1	0	1	0	0	
0	1	1	0	0	0	
0	1	1	1	0	0	
1	0	0	0	0	1	
1	0	0	1	0	0	
1	0	1	0	0	0	
1	0	1	1	0	0	
1	1	0	0	1	1	
1	1	0	1	1	0	
1	1	1	0	1	0	
1	1	1	1	1	0	

Finally we fill in column g by noting that

g = e Nand f

a	b	c	d	e	f	g
0	0	0	0	0	1	1
0	0	0	1	0	0	1
0	0	1	0	0	0	1
0	0	1	1	0	0	1
0	1	0	0	0	1	1
0	1	0	1	0	0	1
0	1	1	0	0	0	1
0	1	1	1	0	0	1
1	0	0	0	0	1	1
1	0	0	1	0	0	1
1	0	1	0	0	0	1
1	0	1	1	0	0	1
1	1	0	0	1	1	0
1	1	0	1	1	0	1
1	1	1	0	1	0	1
1	1	1	1	1	0	1

Binary to Decimal and Decimal to Binary Conversion

When confronted by a binary number such as 110111001, we often need to know what the decimal representation of this number might be. We proceed by noticing that the base is two so we have

$$110111001_2 = 1\cdot 2^0 + 0\cdot 2^1 + 0\cdot 2^2 + 1\cdot 2^3 + 1\cdot 2^4 + 1\cdot 2^5 + 0\cdot 2^6 + 1\cdot 2^7 + 1\cdot 2^8$$
$$= 1 + 0\cdot 2 + 0\cdot 4 + 8 + 16 + 32 + 0\cdot 64 + 128 + 256 = 441_{10}$$

directly from the definition of positional notation. To convert a decimal representation to a binary one is only slightly more complicated. We repeatedly divide the number by two noting the remainders as shown below.

$$\begin{array}{rl}
2\underline{|441} & R \\
2\underline{|220} & 1 \\
2\underline{|110} & 0 \\
2\underline{|55} & 0 \\
2\underline{|27} & 1 \\
2\underline{|13} & 1 \\
2\underline{|6} & 1 \\
2\underline{|3} & 0 \\
2\underline{|1} & 1 \\
0 & 1
\end{array}$$

The binary number is read from the bottom up as 110111001. Notice that the process is not halted until a quotient of **ZERO** (not one) is reached.

Long strings of binary numbers are hard to read and understand; it would help if we could associate the digits of the decimal representation with the binary representation of the individual digits. If we compare 441; the binary equivalent of its digits: 100,100,001; and the binary equivalent of 441: 11011001, we see that the hoped for correspondence does not exist. There is, however a number system for which the equivalence does exist. Consider the base 16 number system called hexadecimal numbers. For base 16 we need 16 digits; we use the usual ten digits followed by A, B, C, D, E, and F.

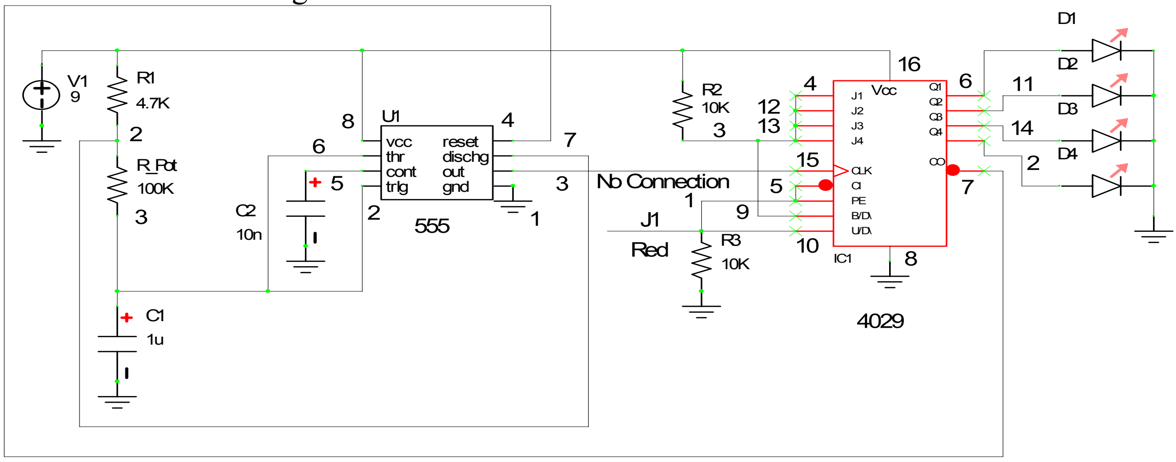
Dec	Hex	Bin		Dec	Hex	Bin		Dec	Hex	Bin	
0	0	0000		16	10	10000		32	20	100000	
1	1	0001		17	11	10001		33	21	100001	
2	2	0010		18	12	10010		34	22	100010	
3	3	0011		19	13	10011		35	23	100011	
4	4	0100		20	14	10100		36	24	100100	
5	5	0101		21	15	10101		37	25	100101	
6	6	0110		22	16	10110		38	26	100110	
7	7	0111		23	17	10111		39	27	100111	
8	8	1000		24	18	11000		40	28	101000	
9	9	1001		25	19	11001		41	29	101001	
10	A	1010		26	1A	11010		42	2A	101010	
11	B	1011		27	1B	11011		43	2B	101011	
12	C	1100		28	1C	11100		44	2C	101100	
13	D	1101		29	1D	11101		45	2D	101101	
14	E	1110		30	1E	11110		46	2E	101110	
15	F	1111		31	1F	11111		47	2F	101111	

Consider the number whose decimal representation is 742; this number has the binary representation 1011100110. The binary digits can be placed in groups of four starting at the **RIGHT** hand side as 10 1110 0110. We can simply replace each group of four binary digits with its hexadecimal equivalent as $2E6_{16}$. W can verify that this is correct by noting that $2E6_{16} = 6 \cdot 16^0 + E \cdot 16^1 + 2 \cdot 16^2 = 6 + 14 \cdot 16 + 2 \cdot 256 = 742_{10}$

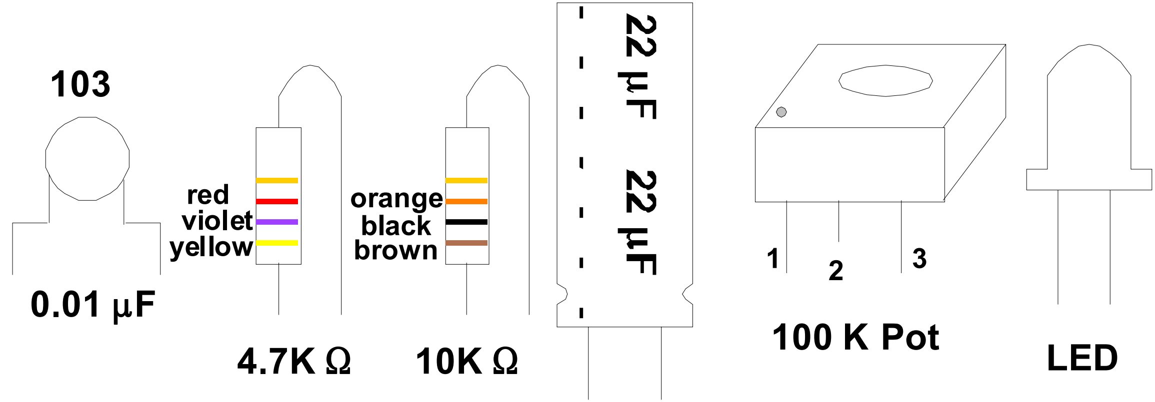
Binary digits are usually referred to as bits (<u>bi</u>nary digi<u>t</u>), and certain sized groups of bits are given special names. An eight-bit number is called a byte, a four byte number is called a nybble (pronounced "nibble"). In general one refers to a binary number containing a specified number of bits as a word. For instance one could have a 32 bit word, a 16 bit word, a 64 bit word etc.

In the Lab

Construct the following circuit.

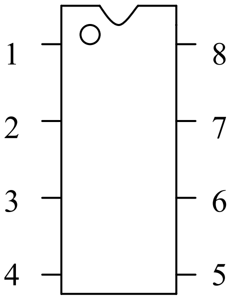


Note that there is NO CONNECTION between pin3 and pin 7 of the 555; between pin 4 of the 555 and 9V; between pin15 of the 4029 and pin 13 of the 4029; or between pin 15 of the 4029 and pin 9 of the 4029. Pins 3, 4, 9, 12, and 13 of the 4029 are connected together to the 10k, and pins 1,5, and 10 of the 4029 are connected together to the other 10k.



Parts Identification

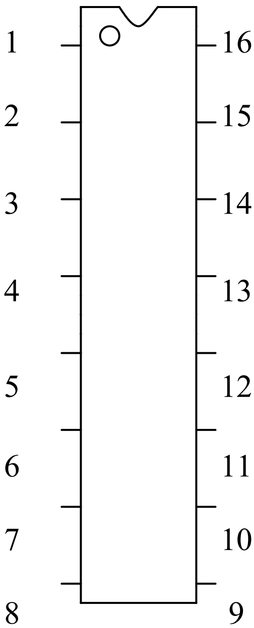
555 pin-out



Pin	Name	Meaning
1	gnd	Ground
2	trig	Trigger
3	out	Output
4	reset	Master Reset (for Flip-flop)
5	cont	Control Voltage
6	thr	Threshold
7	dishch	Discharge
8	vcc	Power Supply

Note that any individual chip may have either a notch, or a circle, or both to indicate the location of pin 1.

4029 pin-out



Pin	Name	Alt Name	Meaning
1	PE	Preset	Preset Outputs to Jn (Pn) values
2	Q4	Q3	Output-MSB
3	J4	P3	Preset-MSB
4	J1	P0	Preset-LSB
5	CI	$\overline{\text{Carry-in}}$	Carry-in active-low
6	Q1	Q0	Output-LSB
7	CO	$\overline{\text{Carry-out}}$	Carry-out active-low
8	gnd		Ground
9	B/D\	Binary / $\overline{\text{BCD}}$	Binary(if hi), BCD(if lo)
10	U/D\	Up / $\overline{\text{Down}}$	Cnt up (if hi) Cnt down (if lo)
11	Q2	Q1	Ouput
12	J2	P2	Preset
13	J3	P1	Preset
14	Q3	Q2	Output
15	CLK	Clock	transitions are counted
16	vcc	Vcc	Power

The outputs may be numbered differently by different manufacturers; some use Q0 – Q3, others use Q1 – Q4. The preset inputs may be numbered P0 – P3 by some manufacturers, J1 – J4 (jam inputs) by others.

Connect the red wire of the 9V battery, then connect the other end of the red wire (J1) whose other end is connected to the 10k connected to pins 1, 5, and 10 to 9V. The four LEDs should now be on continuously. Remove the end J1 from 9V and the counter should count in binary from 1111 (15) to 0000 (0) and stop. You can repeat the cycle by connecting J1 to 9V and disconnecting J1 from 9V. You can vary the speed of the counting by adjusting the pot.

Binary Arithmetic

Binary numbers can be added and subtracted just as decimal numbers are. To add and subtract decimal numbers it is necessary to memorize the addition table, which shows the result of adding any two digits. To add and subtract binary numbers we need only memorize the table listing the result of adding any two binary digits.

	0	1
0	0	1
1	1	10

For example to add the two binary numbers 101100 and 100110 we proceed as follows.

$$\begin{array}{r} {}^{1}10^{1}1^{1}100+ \\ 100110 \\ \hline 1010010 \end{array}$$

Starting at the rightmost column, 0 + 0 = 0 with no carry; 0 + 1 = 1 with no carry; 1 + 1 = 10, so we write the 0 in the sum and carry the one; 1 (from the carry) + 1 +0 = 10, So we write the 0 and carry the one; 1 (from the carry) + 0 + 0 = 1; 1 + 1 = 10, se we write the 0 and carry the 1. In general the addition of two n-bit numbers produces an n+1-bit number (whose first bit may be 1 or 0).

Subtraction is carried out in the usual way but there are many cases where a borrow must be made. For example to subtract 11010111 from 100000001, we proceed as follows

$$\begin{array}{ccccccccccc}
\cancel{1}^{0} & {}^{1}\cancel{0}^{1} & {}^{1}\cancel{0}^{1} & {}^{1}\cancel{0}^{1} & {}^{1}\cancel{0}^{1} & {}^{1}\cancel{0}^{1} & {}^{1}\cancel{0}^{1} & {}^{1}0 & 1 & - \\
 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & \\
\hline
 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 &
\end{array}$$

We proceed as follows. Starting in the rightmost (ones) column, 1 – 1 = 0; in the second (twos) column we have 0 – 1 = ? – since 1 > 0 we must borrow; we look in the next column (fours) and see another 0, so we cannot borrow from it. In fact we must keep searching to the leftmost (256) column before we find a column from which we can borrow. Borrowing from a column is essentially a rewriting of the number in a different format. Before the borrow we write

1(256) + 0(128) + 0(64) + 0(32) + 0(16) + 0(8) + 0(4) + 0(2) + 1;

after the first borrow we write

0(256) + 2(128) + 0(64) + 0(32) + 0(16) + 0(8) + 0(4) + 0(2) + 1.

Then we borrow a 1 from the 128s column and write the number as

0(256) + 1(128) + 2(64) + 0(32) + 0(16) + 0(8) + 0(4) + 0(2) + 1.
Note that all three representations denote the same number (257). We continue in this manner until we have the number represented as
0(256) + 1(128) + 1(64) + 1(32) + 1(16) + 1(8) + 1(4) + 2(2) + 1.
The subtractions are then(starting at the rightmost column), 1 – 1 = 0 (which we have already done); 10 – 1 = 1; 1 – 1 = 0; 1 – 0 = 1;1 – 1 = 0;1 – 0 = 1;1 – 1 = 0; 1 – 1 = 0;
Since all this borrowing can become tedious and error-prone, an algorithm was developed to enable the carrying out of a subtraction with steps that involve only adding and logical complementation. Recall that the logical compliment of 1 is 0, and the logical compliment of 0 is one. We can extend the notion of the logical compliment to binary numbers by taking the original number and changing all the 1s to 0s, and all the 0s to 1s. The logical compliment of a binary number is called its ones compliment. We are now in a position to define our subtraction of A – B: $A - B = A + B^{2C}$ ignoring any carry out that may be generated by the addition. B^{2C} is called the twos complement and is found by adding 1 to the ones complement of B, i.e. $B^{2C} = \overline{B} + 1$. In order for this algorithm to work A and B MUST be expressed using the same number of bits. For our example we have A = 100000001 and B = 011010111; so $\overline{B}$ = 100101000, and B^{2C}=100101000+1 = 100101001; and 100000001-011010111 = 100000001 + 100101001 = 00101010 as before. (we ignore the carry out.)

$$ {}^{1}10000000{}^{1}1+ $$

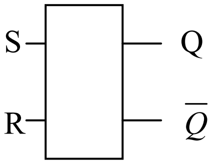
$$ \underline{1001\ 0100\ 1} $$

$$ 10001\ 010\ 1\ 0 $$

Binary number systems that permit the representation of negative numbers are called twos complement numbersystems.

Sequential Logic

The logic we have studied so far is called combinational logic – the output(s) depend only on the state of the current input(s). There is another family of logic, which we have encountered in the 555 timer, called sequential logic. Sequential logic is the study of devices called flip-flops, whose output depends not only on the current state of the inputs (hi or lo) but on PAST states of the inputs. We shall investigate only the two simplest flip-flops of the many that are available.
The simplest flip-flop is called an S-R (for Set - Reset) flip-flop. To understand this device, show below, we must construct its excitation table. (In digital jargon, set means to make something 1, and reset means make something 0.)

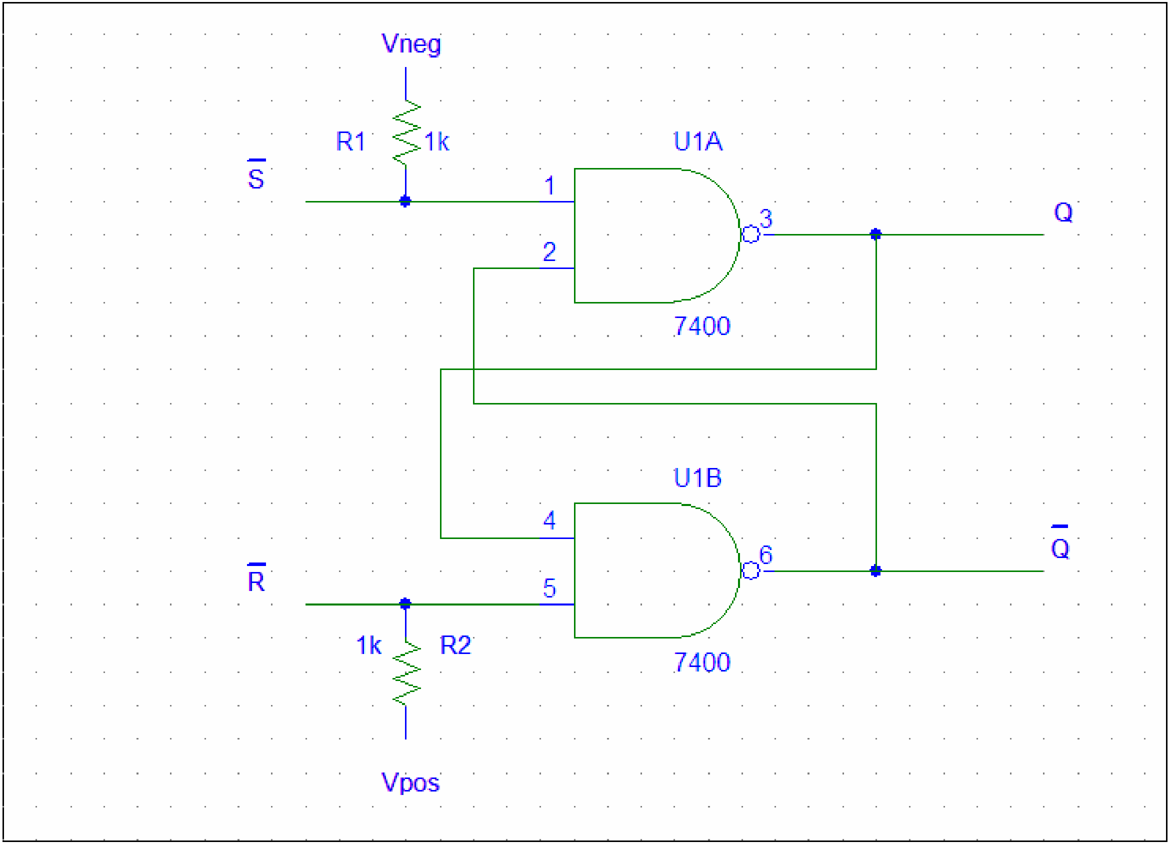


S R Flip-Flop

S	R	Q
0	0	No Change
1	0	1
0	1	0
1	1	Undefined

The excitation table is interpreted as follows. If the S input becomes active (goes hi – becomes true) then Q becomes active (goes hi – becomes true), and stays that way even if S becomes inactive (goes lo – becomes false). If the R input becomes active (goes hi – becomes true) then Q becomes inactive (goes lo – becomes false), and stays that way even if R becomes inactive (goes lo – becomes false). The fact that the state where BOTH inputs are low does not uniquely determine the state of the flip-flop, means that when the device "wakes up" i.e. , when power is first applied to the device, we cannot know for sure what the state of Q will be. The state of $\overline{Q}$ is always the opposite from that of Q.

An S R Flip-flop can be constructed out of a pair of cross-coupled nand gates as shown below. (Note that the inputs are named, $\overline{S}$ and $\overline{R}$, not S and R: this indicates that the inputs must become 0 to do their jobs. Such inputs are called active low and are notated with an overbar.)


Vneg
R1
1k
U1A
S̄
1
3
Q
2
7400
U1B
4
6
Q̄
R̄
5
1k
R2
7400
Vpos

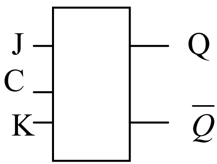
There are two equally likely possibilities:
case 1 – Q wakes up 1,
case 2 – Q wakes up 0.

Case 1. Since Q is 1, $\overline{Q}$ is 0 and the inputs to the top nand gate are 1 and 0, which keeps the output of the top nand gate 1. Since Q is 1, the inputs to the bottom nand gate are 1 and 1 which keeps the output of the bottom nand gate 0.

Case 2. Since Q is 0, $\overline{Q}$ is 1 and the inputs to the top nand gate are 1 and 1, which keeps the output of the top nand gate 0. Since Q is 0, the inputs to the bottom nand gate are 0 and 1 which keeps the output of the bottom nand gate 1.

Now whenever Q is 1, then $\overline{Q}$ is 0 and the inputs to the top nand gate are either 0 and 1, or 0 and 0, either of which makes the output of the <u>top</u> nand gate 1. This means that if Q is 1, the current state of $\overline{S}$ does not affect the output of the top nand gate, i.e. , once the flip-flop is set, further manipulation of its $\overline{S}$ input does not matter. On the other hand if Q is 1, then $\overline{Q}$ is 0 and the inputs to the <u>bottom</u> nand gate are 1 and 1, which keeps the output of the bottom nand gate 0 only as long as $\overline{R}$ remains 1, if $\overline{R}$ goes low (0), then the inputs to the bottom nand gate become 1 and 0 which makes the output of the bottom nand gate 1, which makes the inputs to the top nand gate 1 and 0, which makes the output of the top nand gate (Q) 0. The inputs to the bottom nand gate are now either 0 and 1 (if $\overline{R}$ is 1) or 0 and 0 (if $\overline{R}$ is 0), in either case the output of the bottom gate stays 1. Once the flip-flop is reset, further activity on the $\overline{R}$ input does not matter. If the flip-flop is reset (Q = 0 and $\overline{Q}$ = 1) and $\overline{S}$ goes low, the inputs to the top nand gate are 0 and 1, which changes the output of the top nand gate to 1, which changes the inputs to the bottom nand gate to 1 and 1, which changes the output of the bottom nand gate to 0.

Another simple flip-flop is the J K flip-flop shown below. Note that the set input is called J and the reset input is called K. The principle difference is that for a J K flip-flop activating both inputs at once toggles the output at the next clock input. (If Q was 1 it becomes 0, if Q was 0 it becomes 1 at the next active transition of the clock C.) Note that the J K flip-flop is a clocked device the outputs only change on the active transition of the clock pulse.



J	K	Q
0	0	No Change
1	0	1
0	1	0
1	1	$\overline{Q}$

J K Flip-Flop

Computers and Programming I: Basic

The first four pages were excerpted from the SmallBASIC documentation file

…\SmallBASIC\doc\ guide.html

1. Introduction

1.1 Welcome to SmallBASIC

SmallBASIC (SB) is a simple computer language, featuring a clean interface, strong mathematics and string library. We feel it is an ideal tool for experimenting with simple algorithms, for having fun.

1.1.1 About BASIC

BASIC is a very simple language and it is a perfect tool for calculations or utilities. Its name stands for (B)eginners (A)ll-purpose (S)ymbolic (I)nstruction (C)ode. It was developed by John Kemeny and Thomas

Kurtz at Dartmouth College during the middle of 1960, and was one of the most popular languages for several decades.

However, at the last decades it was upgraded to survive on the new programming environments. It was modernized and that was hard required.

In the first upgrade, BASIC was transformed to a structured language. As far, as I known, the first structured BASIC was the QuickBASIC (QB), a Microsoft product. Several structured dialects was followed from other companies.

In the second upgrade, BASIC was transformed to an (almost) object-oriented language. As far, as I known, the first OO BASIC was the VisualBASIC (VB), a Microsoft product. In that stage BASIC was become very problematic, since, Microsoft was introduced ObjectPascal and C++ technologies in a language with very different design and purpose of existance!

Anyway, we strongly disagree with the "new" feautures and the way that are implemented in VB. Every language created for specified purposes, BASIC for beginners, C for low-level programming, Prolog for

AI, etc. VB it is not object-oriented nor a simple language (anymore), but it is a bad designed mix of other languages.

1.1.2 About SmallBASIC

SmallBASIC was created by Nicholas Christopoulos in May of 2000, to be used as an advanced calculator for his Palm IIIx handheld device. In Jan of 2001, SB moved to the web as an GPL project.

Because SB was designed for that small device (Palm IIIx), and because was small compared to desktop-computer BASICs, it takes the prefix 'Small'.

SB is a structured version of BASIC and includes a lot of new feautures such matrices, algebra functions, powerfull string library, etc. A lot of its feautures does not exists in the most languages, but on the other hand, SB does not supports GUI and other feautures that are common in today languages.

1.1.2.1 Purpose

BASIC is easy to learn and simple to use, and this is the spirit of SB. Instead of other BASIC versions, as VB, our version intent to sucrifice everything in the altar of simplicity.

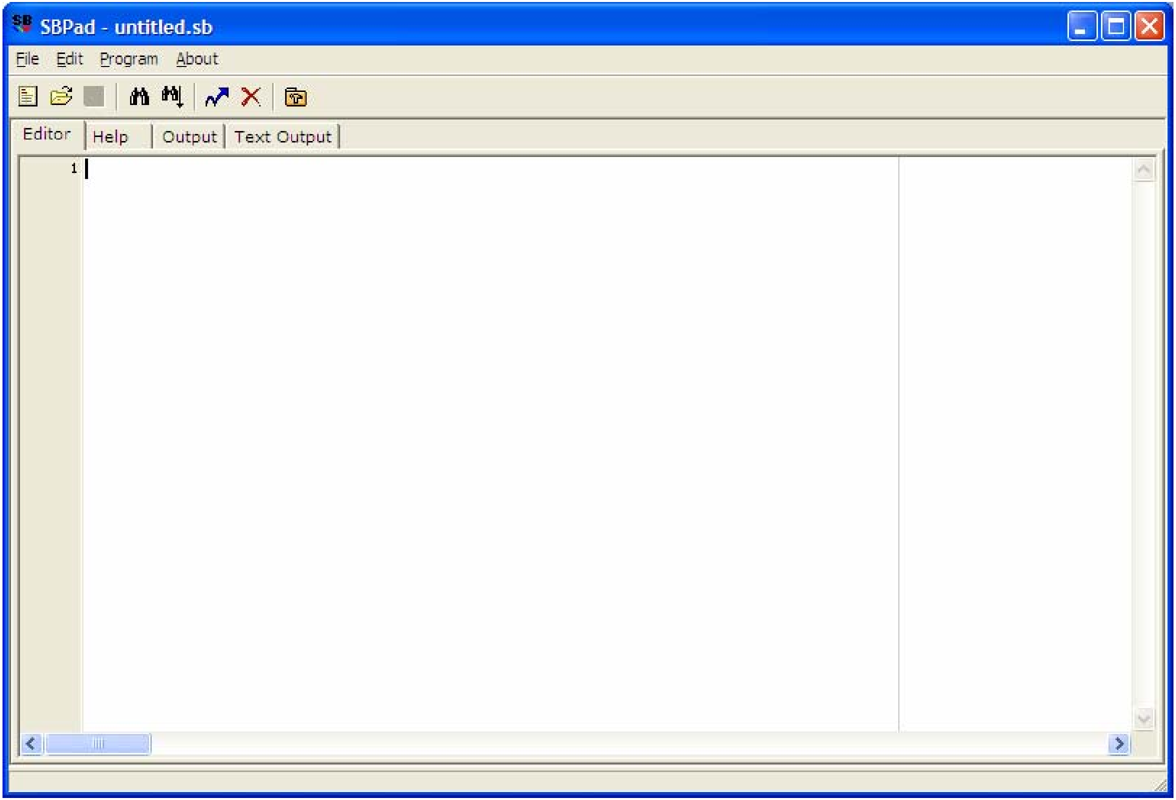
The world is full of languages, SB does not offers something new, but intents to offer what is lost in our days. A simple tool for easy to write programs, an easy way to do some maths and build some scripts.

Our priorities are to build

- An extremly easy learned language.
- An extremly easy to use language.
- An ideal tool for experimenting on programming.
- An excellent tool for mathematics.
- An excellent tool for shell-scripts.

We shall not make any effort to study the art of computer programming or the Basic language in any detail. We wish only to introduce the subject of computer programming in enough depth to cover four main concepts of the subject: assignment, I/O (Input/Output), loops, and conditional statements (If statements). We shall study enough Basic (in SmallBASIC) to enable us to successfully apply the Basic-Stamp micro-controller in the next chapter.

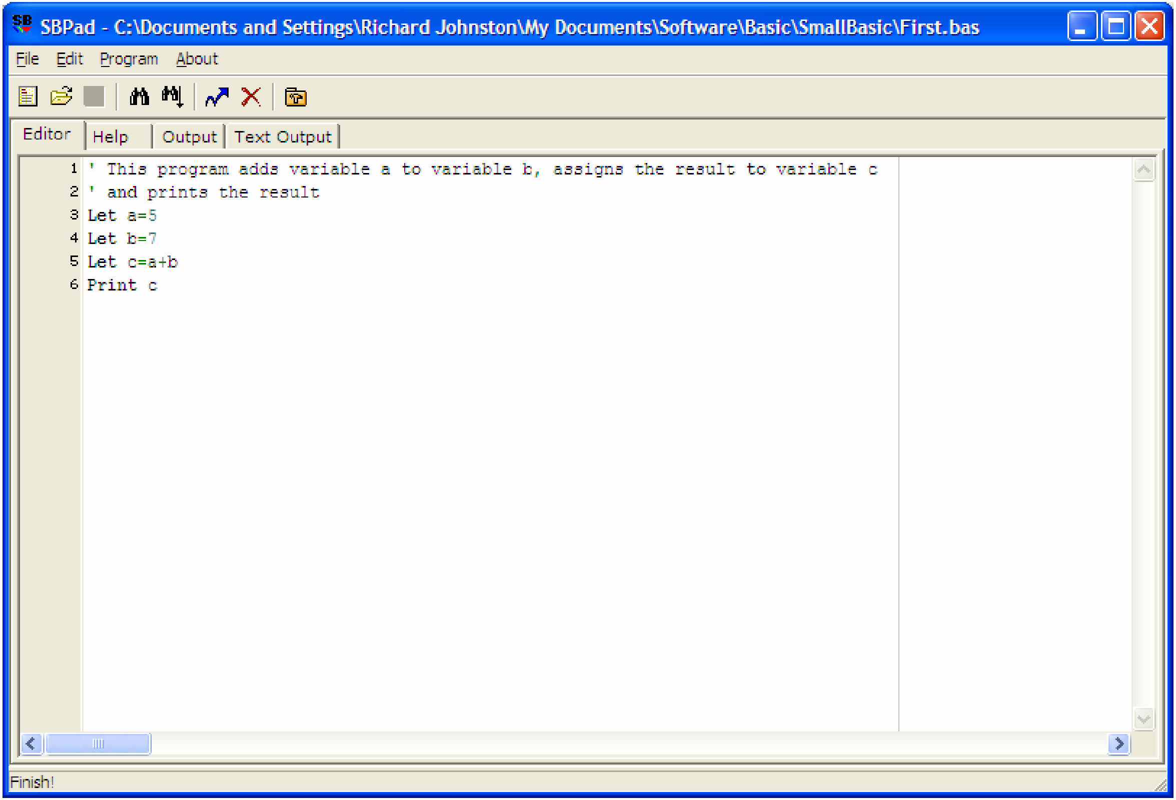
You must first obtain the file  smallbasic-0.9.5-wingui-setup.exe either from the Black-board site for this course, the cd-rom for this course (which may be obtained from your instructor) or from http://source forge.net, and execute the file to install the software. It is a good idea to locate the file …\SmallBASIC\  Sbpad.exe and create a shortcut to it on the desk-top. When you double-click on the shortcut you will see the following screen.



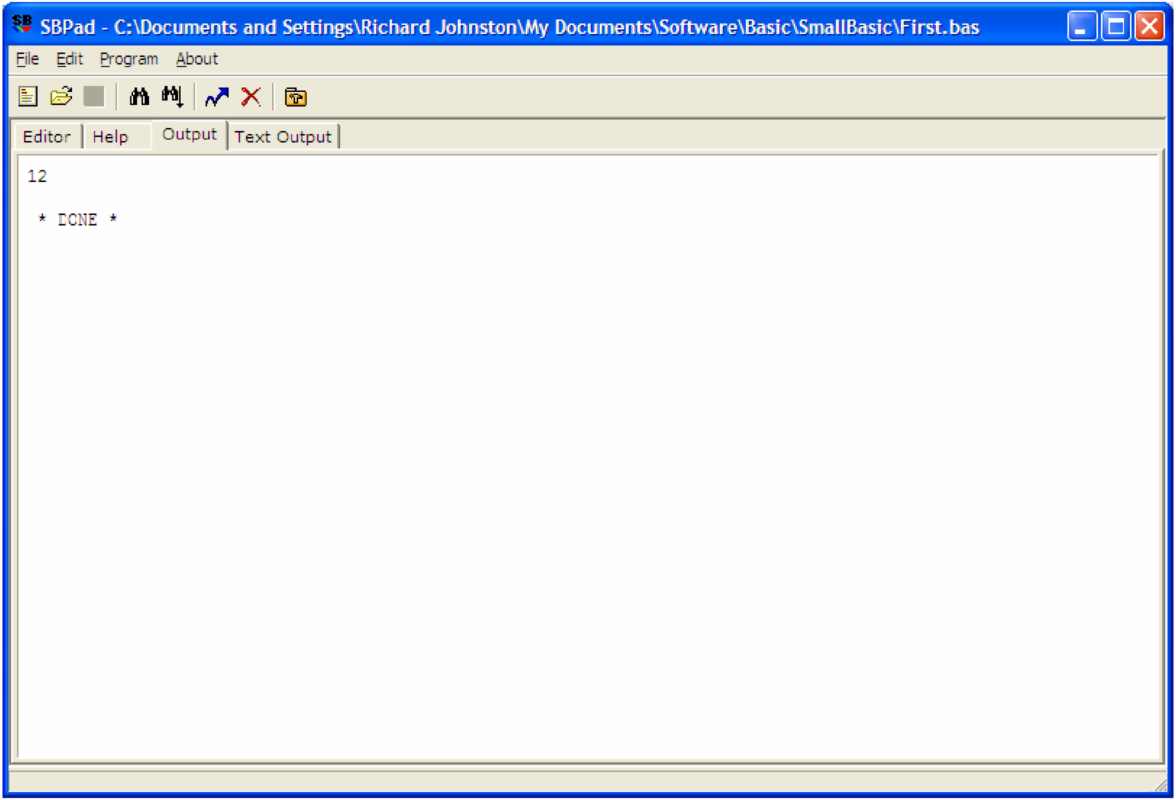
The area on the left is for line numbers that the editor automatically inserts and keeps updated.

A First Program

Our first program illustrates assignment and output, by assigning the value 5 to variable a, assigning the value 7 to variable b, assigning the sum of a and b to variable c, then printing the result.

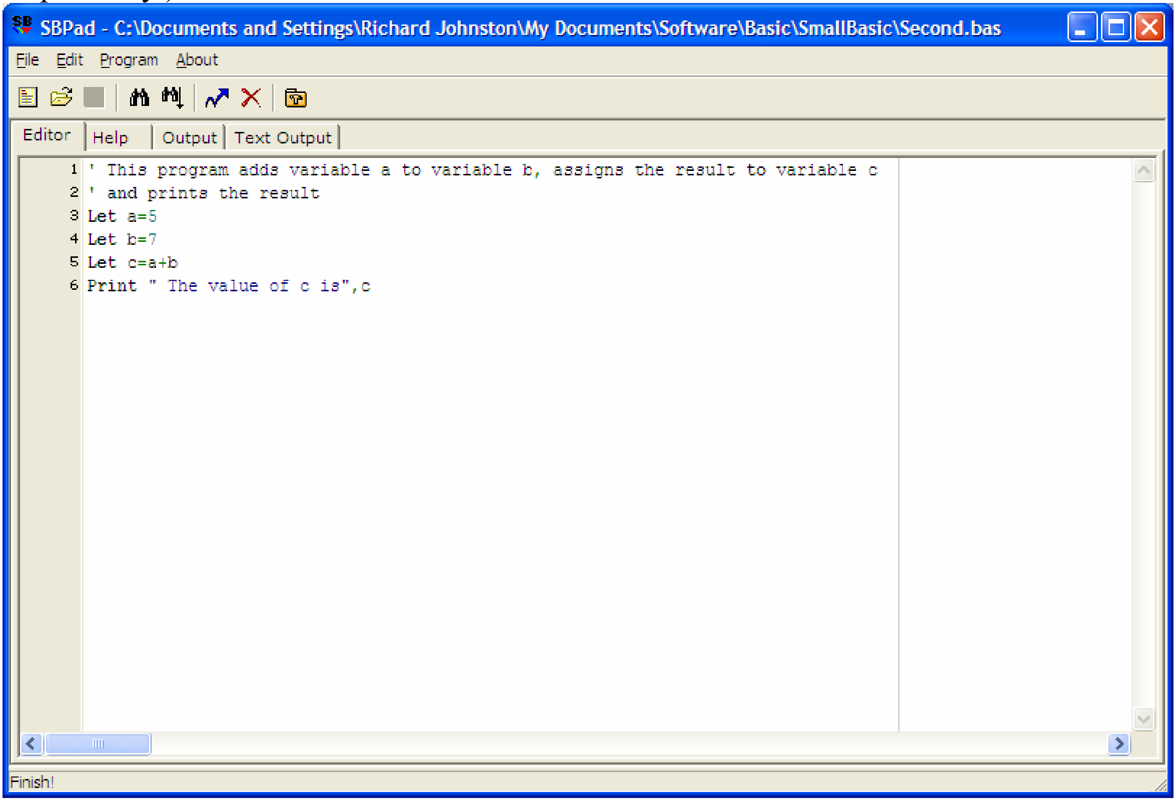


If we click Program, Run, we see the following screen.

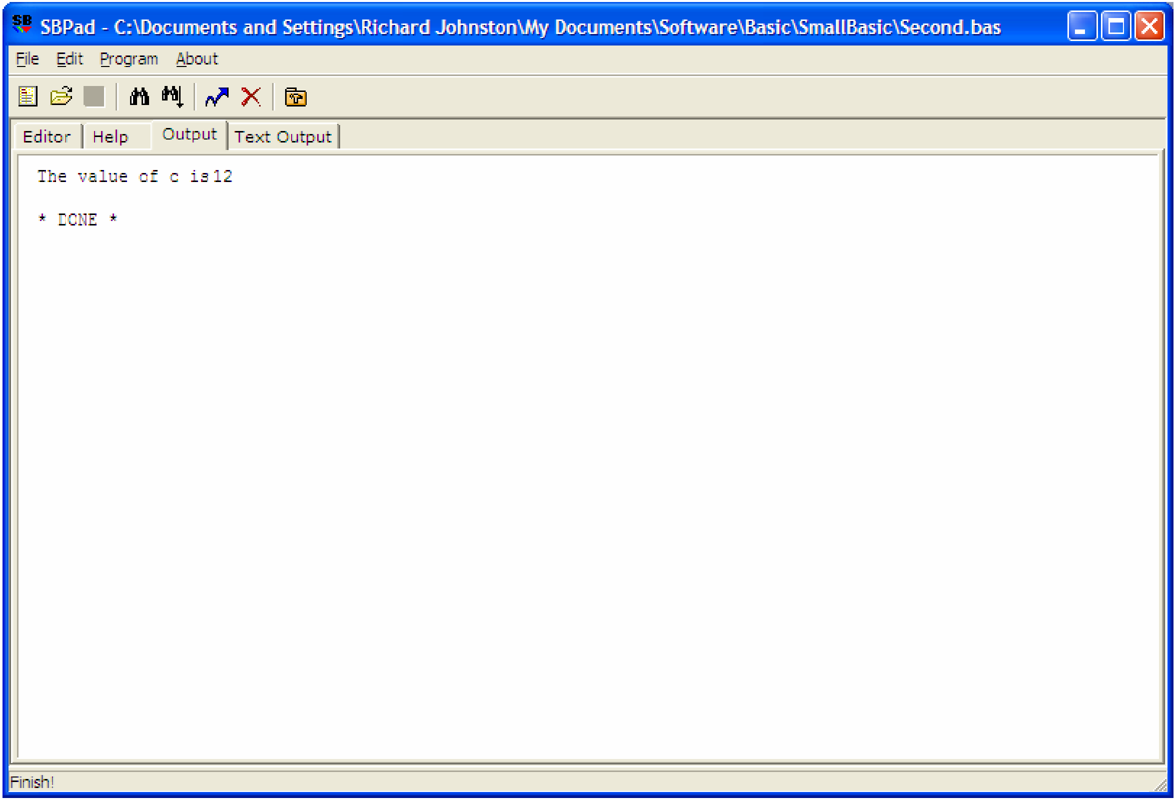


The 12 is the value of variable c as specified in the print statement. We can greatly improve things, by clicking the editor tab and modifying the print statement as follows.

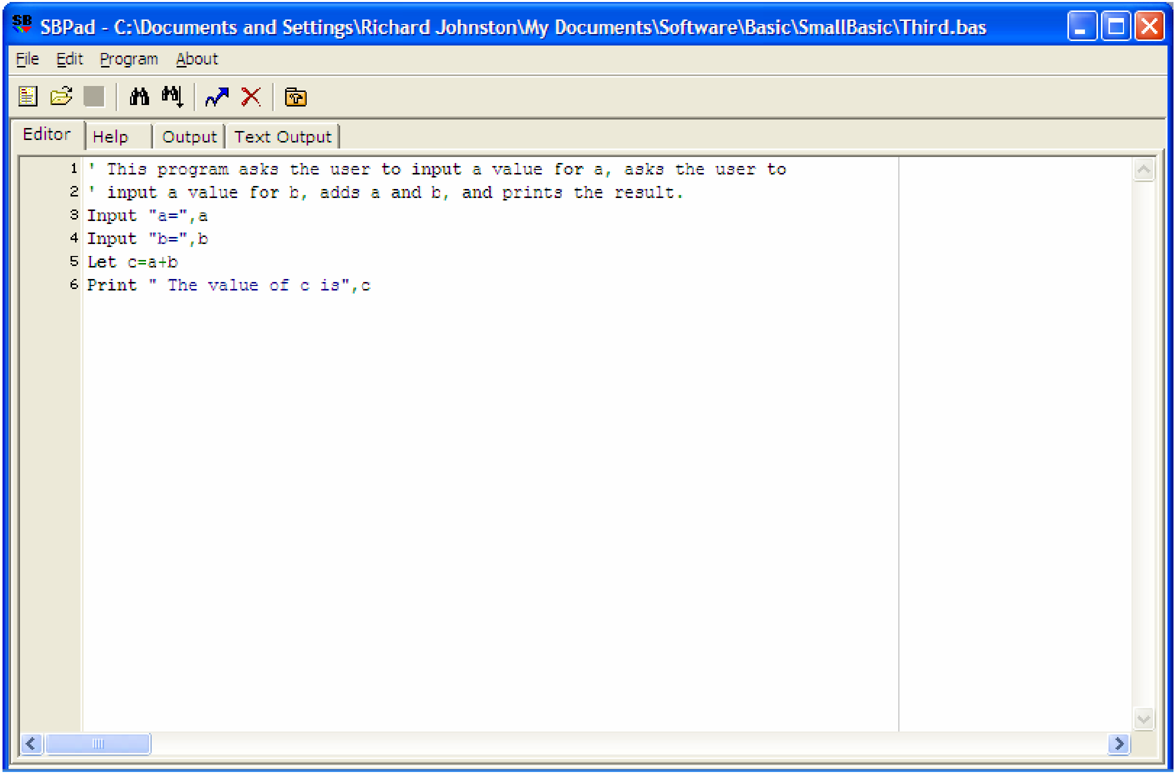
(We move between the program and its output by clicking the Editor and Output tabs respectively.)

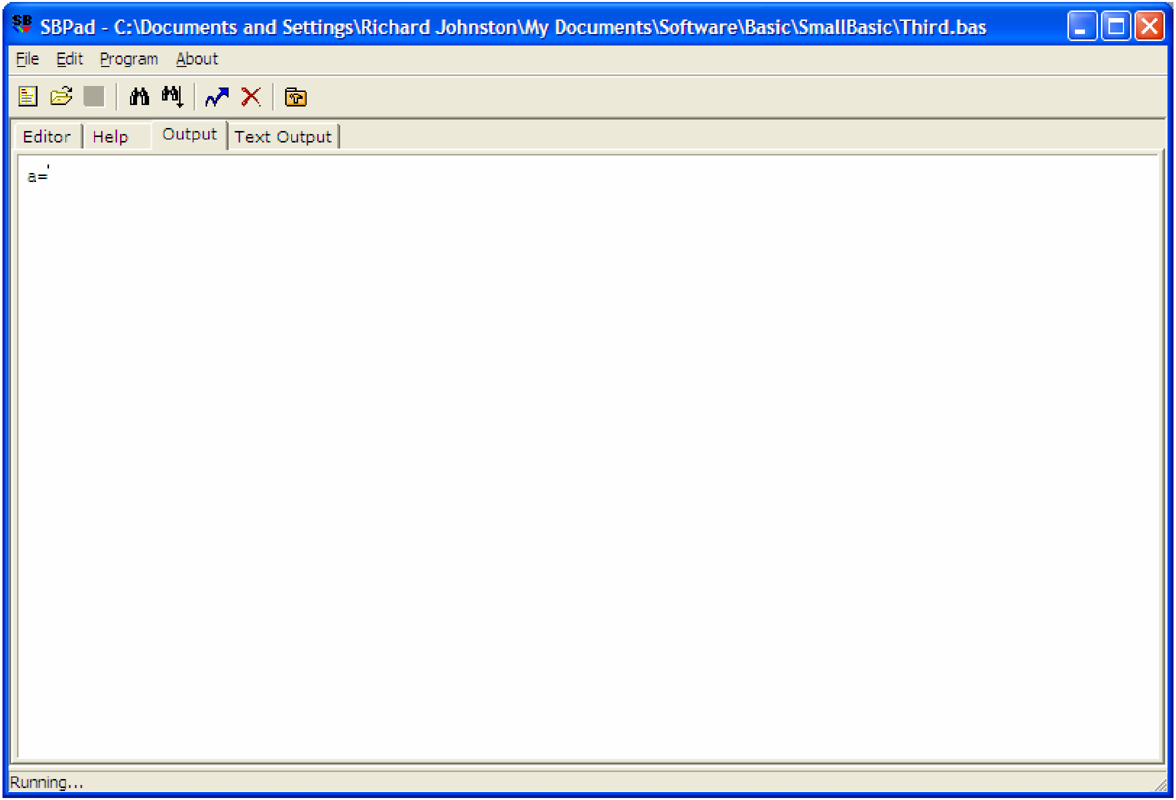


The output of this second program is shown below.

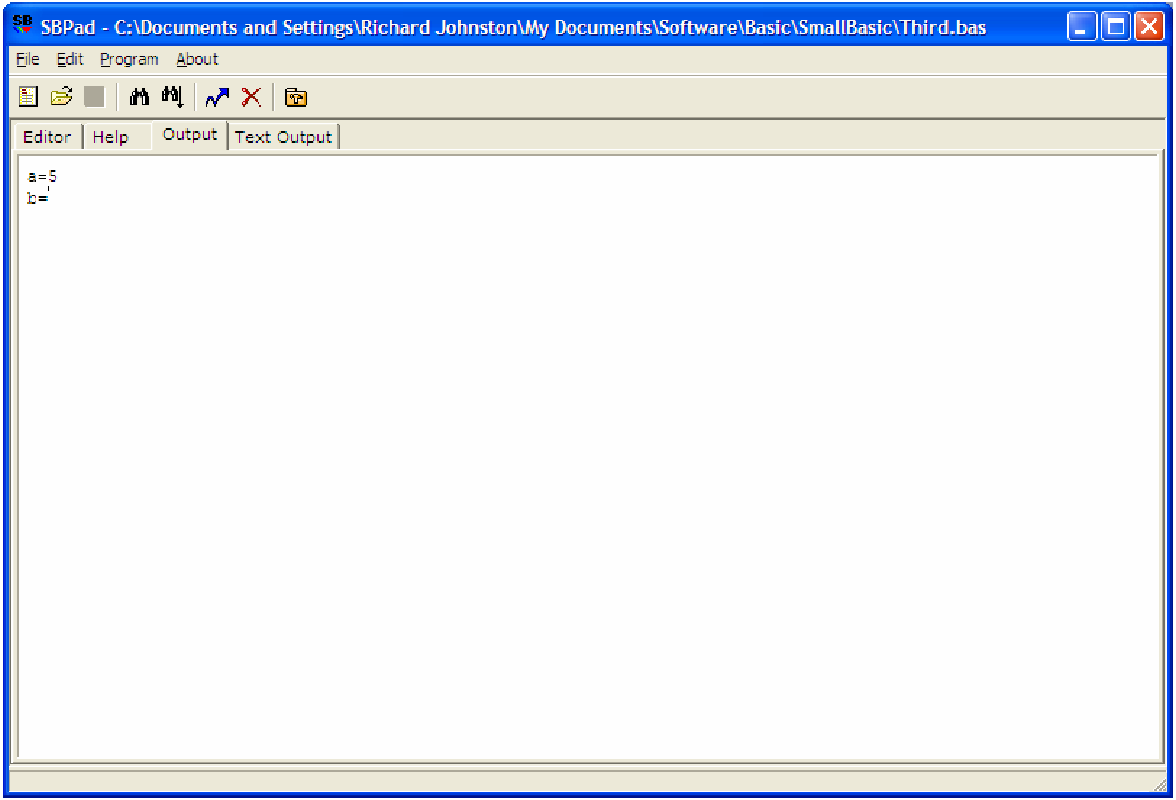


Of course, we rarely would write a program to add two *specific* numbers together, we would write a program to add any two numbers together. To do this we need a way to ask the user for input. Study the third program and the various output screens it produces.

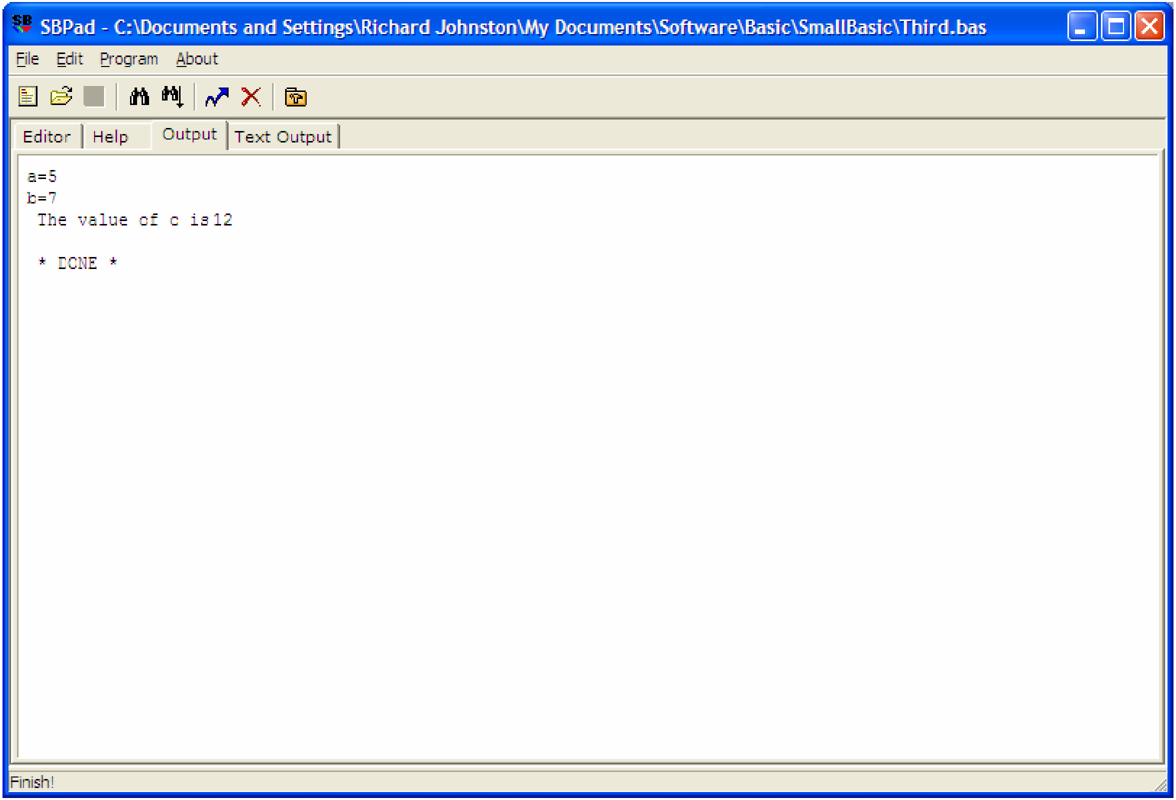

SBPad - C:\Documents and Settings\Richard Johnston\My Documents\Software\Basic\SmallBasic\Third.bas
File Edit Program About
Editor Help Output Text Output
' This program asks the user to input a value for a, asks the user to
' input a value for b, adds a and b, and prints the result.
Input "a=",a
Input "b=",b
Let c=a+b
Print " The value of c is",c



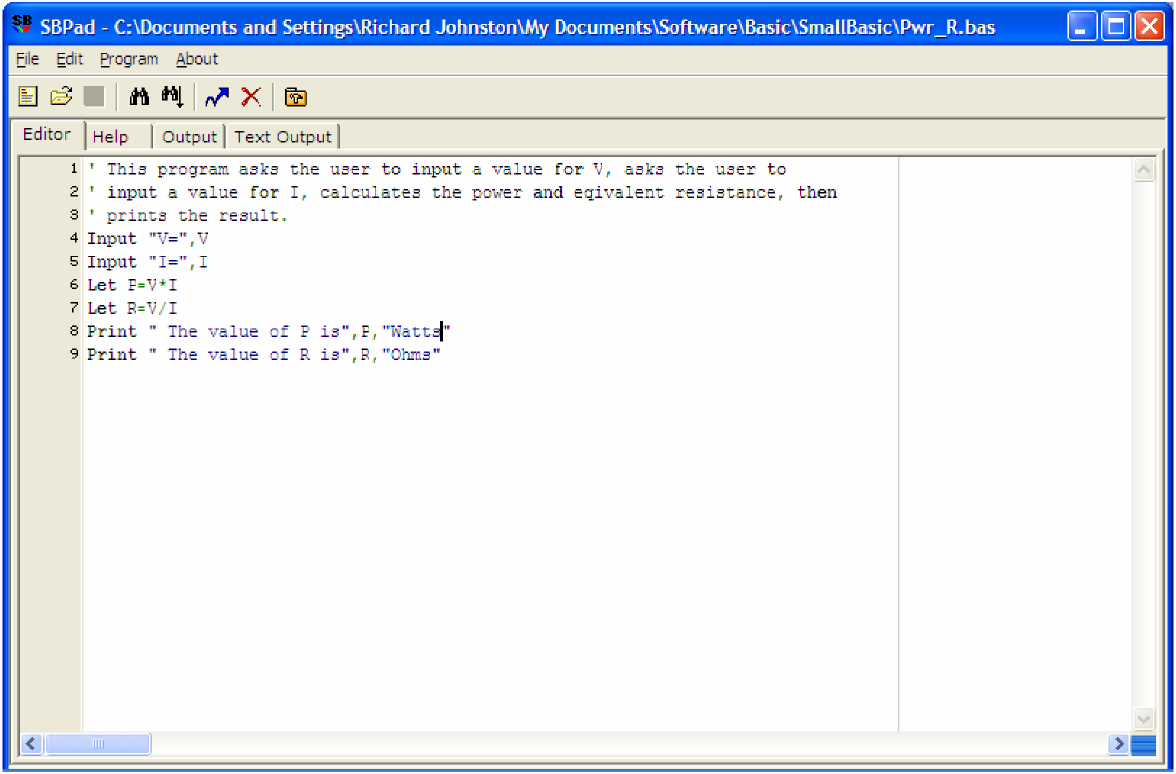
This screen asks us for a value for a so we type, for example, 5 (followed by the enter key), which produces the following screen.

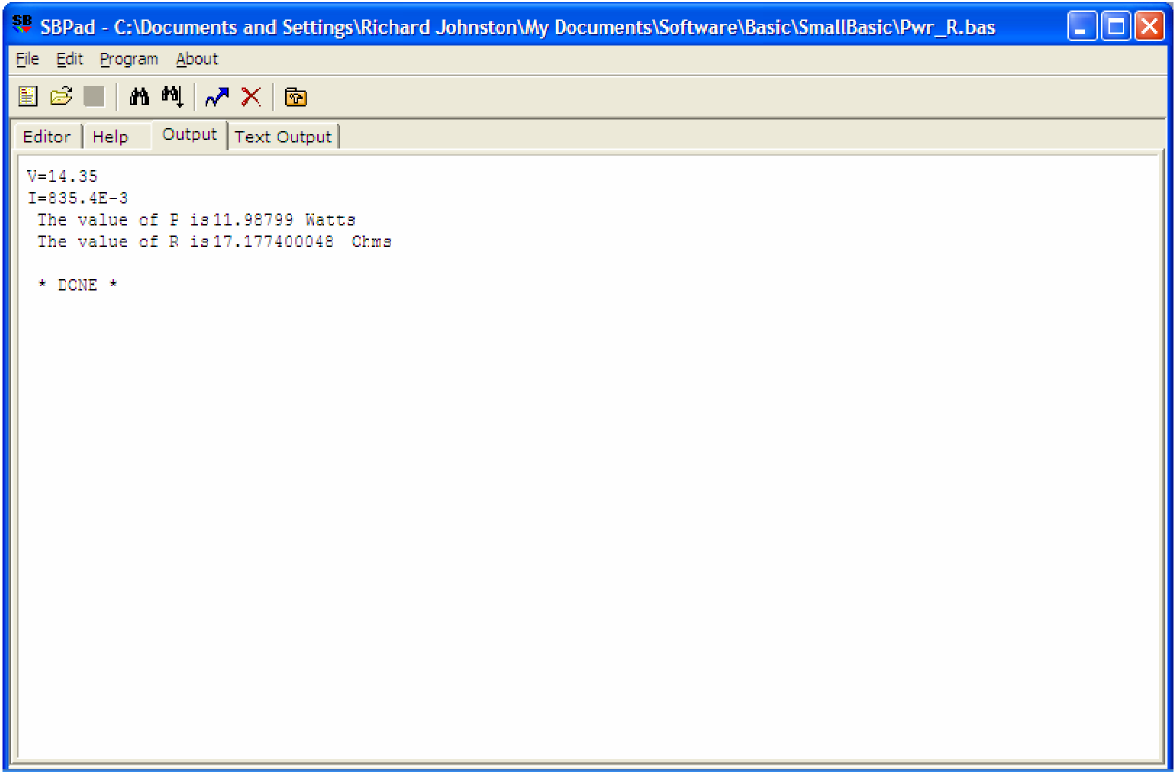


This screen asks us for a value for b so we type, for example, 7 (followed by the enter key), which produces the following screen.



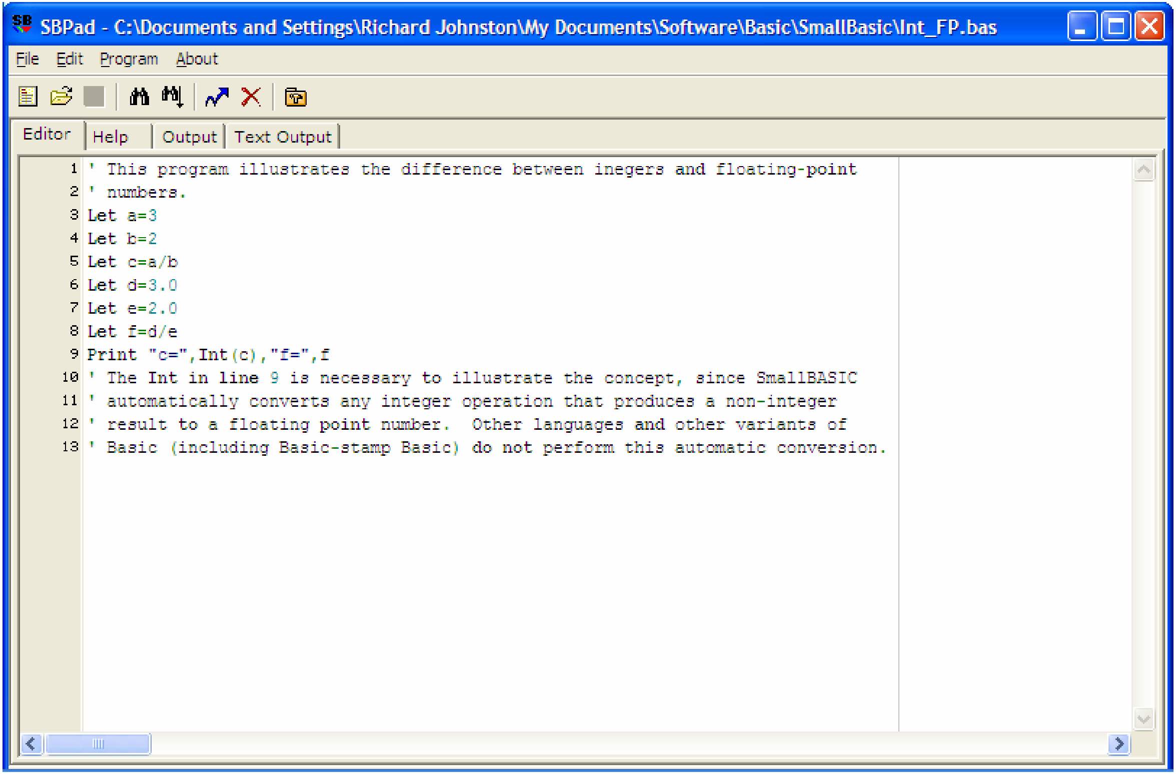
The following program illustrates the calculation of power and equivalent resistance from measurements of voltage and current. Let us assume that the measurements yield V=14.35V , and I=835.4mA.

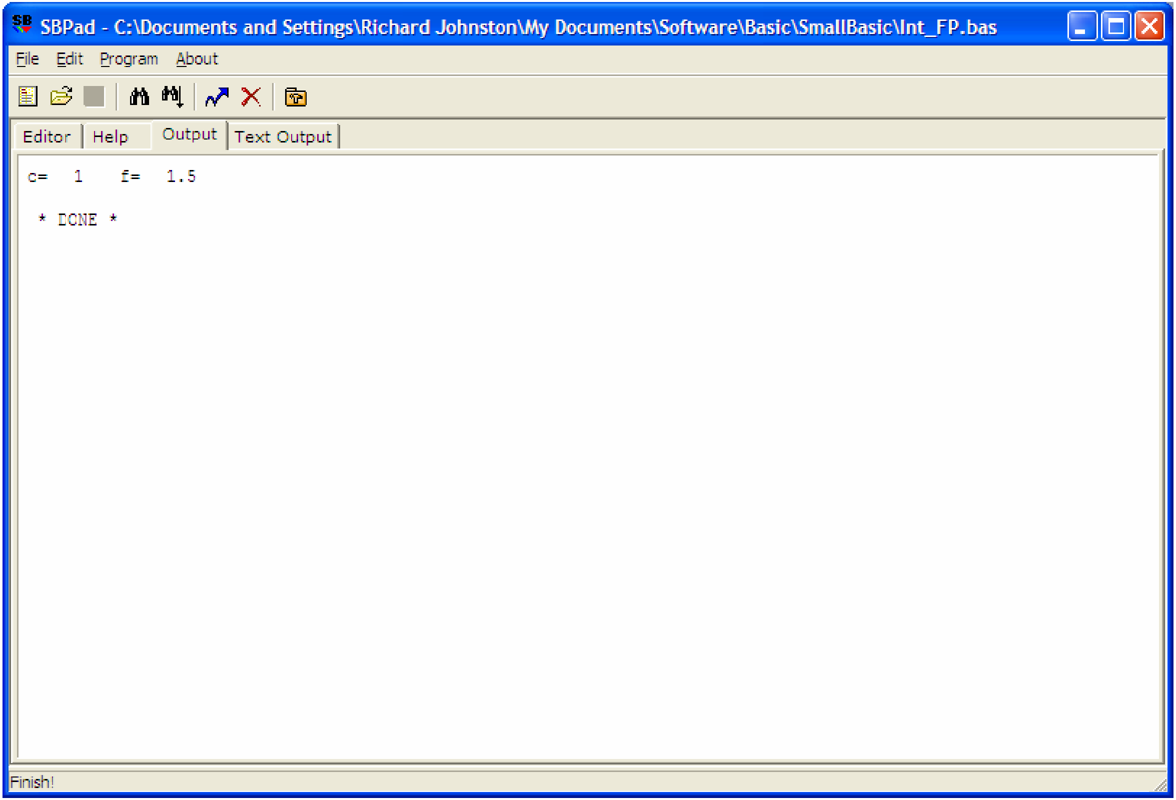

SBPad - C:\Documents and Settings\Richard Johnston\My Documents\Software\Basic\SmallBasic\Pwr_R.bas
File
Edit
Program
About
Editor
Help
Output
Text Output
' This program asks the user to input a value for V, asks the user to
' input a value for I, calculates the power and eqivalent resistance, then
' prints the result.
Input "V=",V
Input "I=",I
Let P=V*I
Let R=V/I
Print " The value of P is",P,"Watts"
Print " The value of R is",R,"Ohms"



Note the use of 835.4E-3 to represent $835.4 \cdot 10^{-3}$. This is called exponential notation. Of course 835.4E3 would represent $835.4 \cdot 10^{3}$.

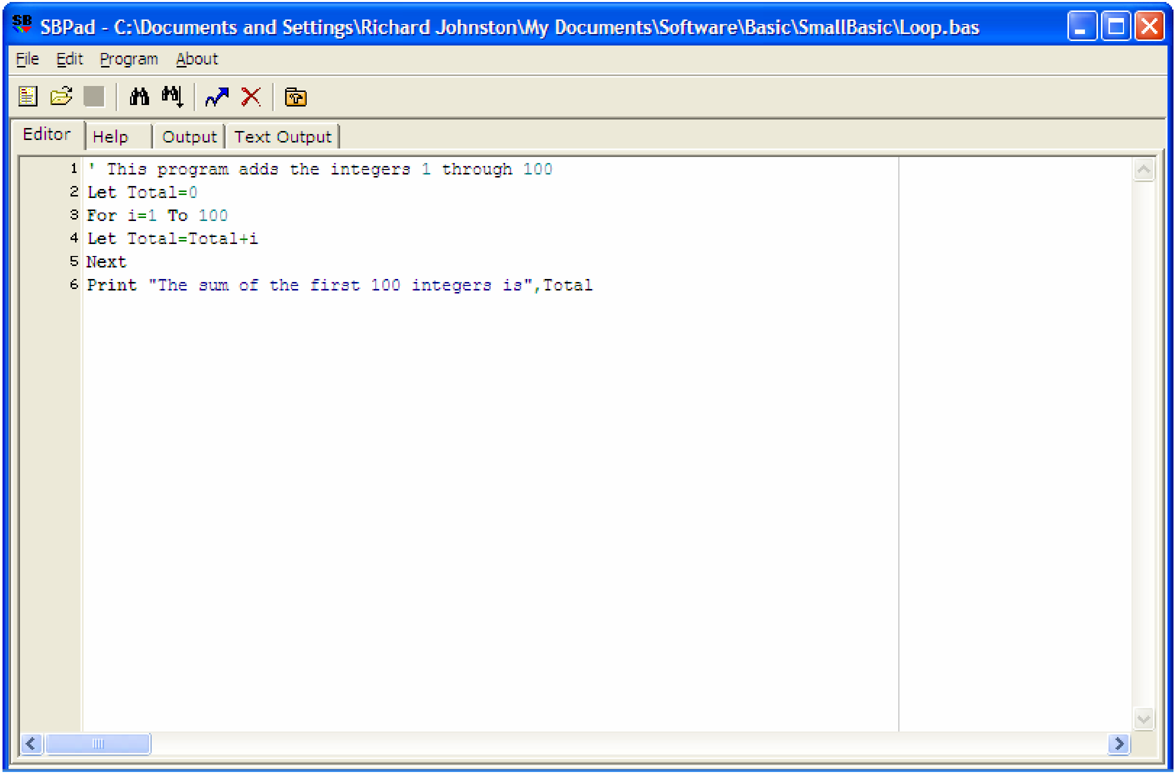
At this point we pause to consider the two data structures in basic with which we shall be concerned: integers and floating point numbers. Integers are counting numbers like 1, 2, 3, -1, -2, -3, etc. Floating point numbers are numbers that have fractional parts like 3.4, 5.2, 3.14159, etc. It is important to understand that Basic (and indeed all the computer languages that you will study) treats the number 3 differently from the number 3.0. The number 3 is an integer, while the number 3.0 is a floating point number. When most variants of Basic (and most other computer programming languages) do arithmetic operations on integers they always provide an answer that is an integer. For addition, subtraction, and multiplication this poses no problems, but for division the result of 3/2 is different from the result of 3.0/2.0. 3/2 = 1, but 3.0/2.0 = 1.5

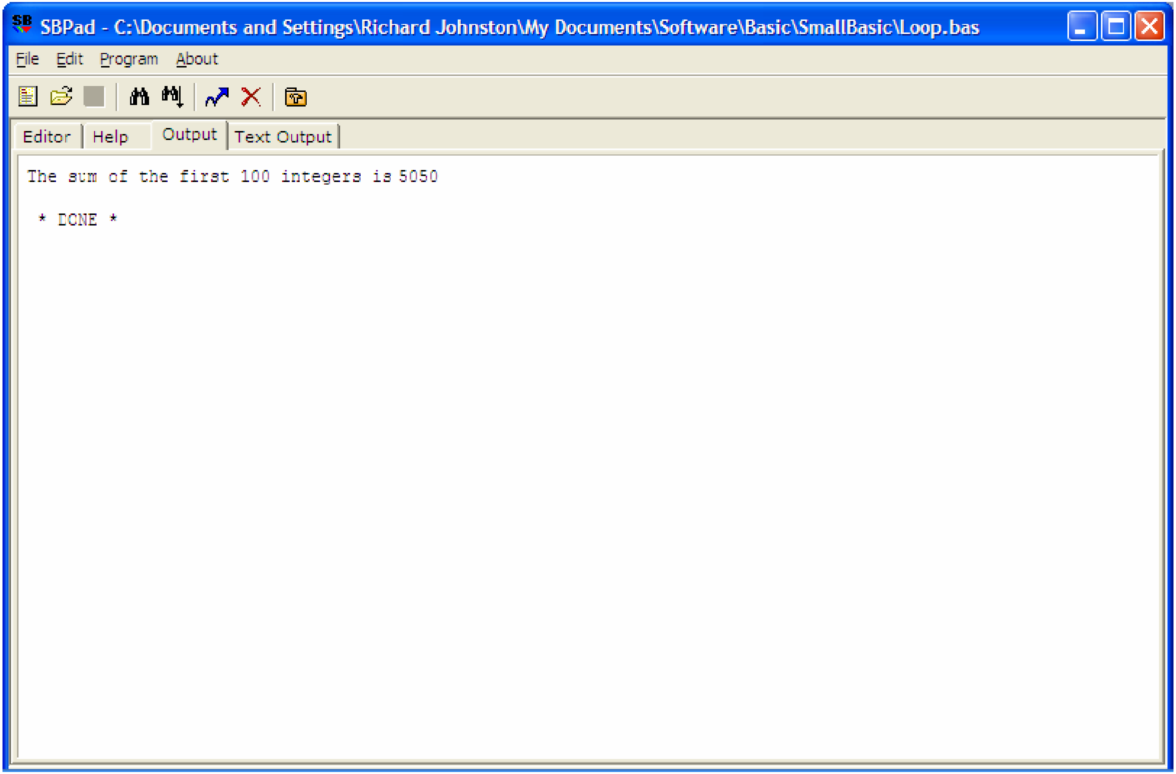

SBPad - C:\Documents and Settings\Richard Johnston\My Documents\Software\Basic\SmallBasic\Int_FP.bas
File Edit Program About
Editor
Help
Output
Text Output
1 ' This program illustrates the difference between inegers and floating-point
2 ' numbers.
3 Let a=3
4 Let b=2
5 Let c=a/b
6 Let d=3.0
7 Let e=2.0
8 Let f=d/e
9 Print "c=",Int(c),"f=",f
10 ' The Int in line 9 is necessary to illustrate the concept, since SmallBASIC
11 ' automatically converts any integer operation that produces a non-integer
12 ' result to a floating point number. Other languages and other variants of
13 ' Basic (including Basic-stamp Basic) do not perform this automatic conversion.


SBPad - C:\Documents and Settings\Richard Johnston\My Documents\Software\Basic\SmallBasic\Int_FP.bas
File Edit Program About
Editor Help Output Text Output
c= 1 f= 1.5
* DONE *
Finish!

Loops

Computer programs are at their best when automating long, tedious calculations. Let us consider adding together the first 100 integers. (This ambiguous statement could mean add the digits 0 through 99 of add the digits 1 through 100 – it all depends on whether you count 1 as the first integer or 0 as the first integer. We shall add the integers 1 through 100.) The calculation consists of an initialization (Sum=0) and an iteration (the For Next loop).


SBPad - C:\Documents and Settings\Richard Johnston\My Documents\Software\Basic\SmallBasic\Loop.bas
File Edit Program About
Editor Help Output Text Output
' This program adds the integers 1 through 100
Let Total=0
For i=1 To 100
Let Total=Total+i
Next
Print "The sum of the first 100 integers is",Total



Special attention should be given to line 4. Note that an assignment statement is NOT a statement of equality – it is an instruction to do something! Namely, to assign to the variable on the left side of the equals sign, the value of whatever appears on the right side of the equals sign.

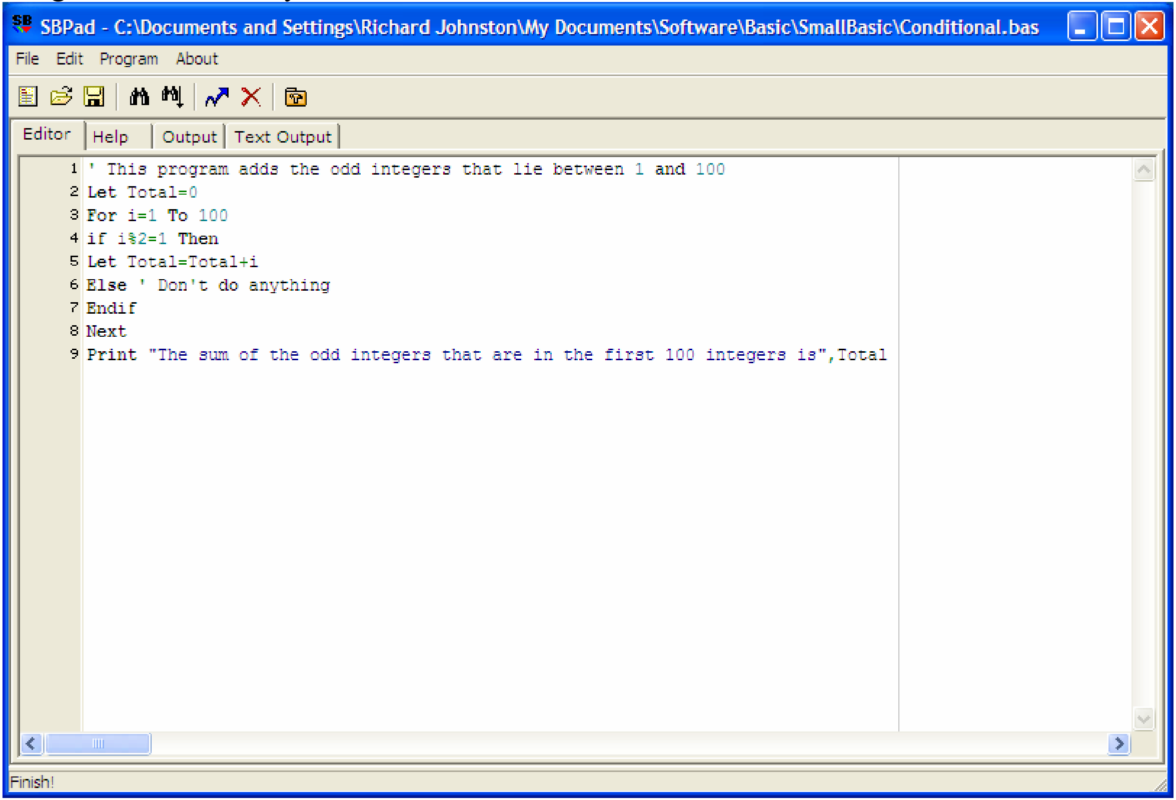
Conditional Statements

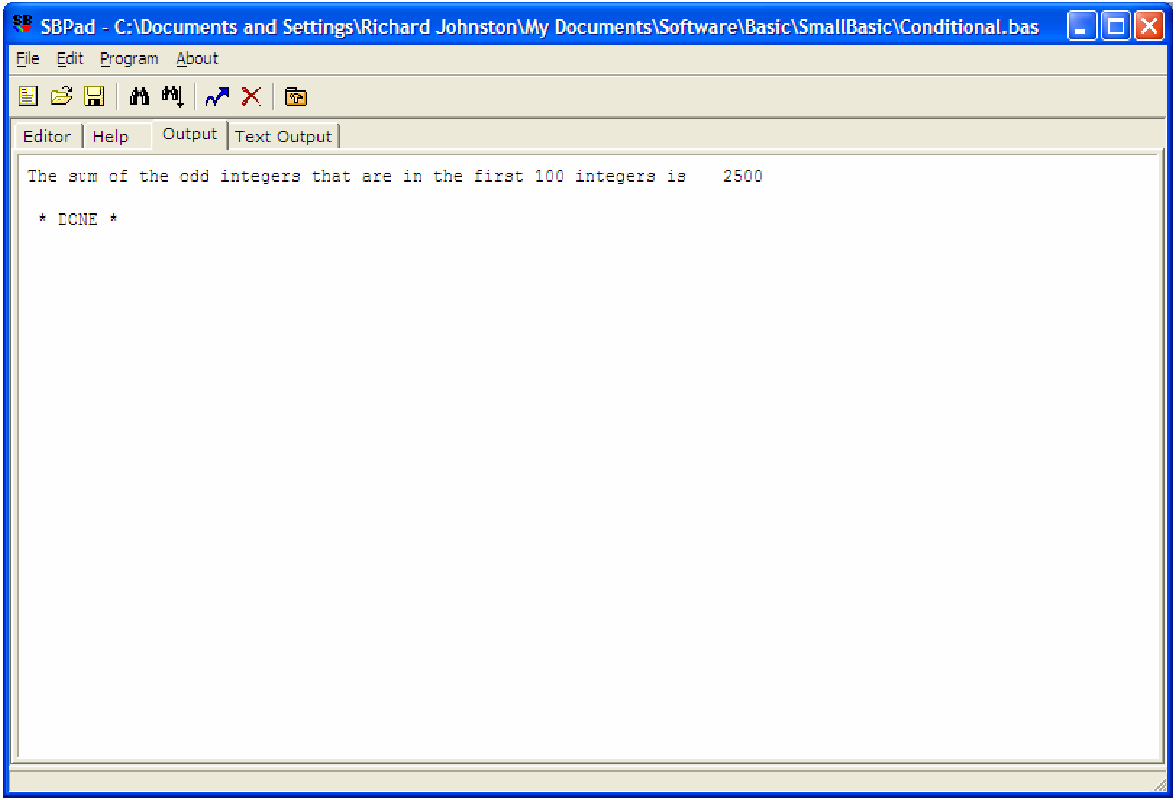
It is often necessary to do something different depending on the truth of some condition. For example, to add together only the odd integers that occur in the first 100 integers we could modify our program to only add an integer i to Total if i is odd. The general form of the conditional statement (If statement) is

If *condition* Then
Statement(s)
Else
Statement(s)
Endif

where condition is a statement that can be either true or false, for example:
a<b
a>b
a=b
etc.
For our purpose we need a statement that is true if i is odd, and false if i is even. Consider the % operator (the Mod statement).

a%b gives the remainder of the integer division of a by b. a%2 is the remainder of the integer division of a by 2, so a%2=0 if a is even and a%2=1 if a is odd.




SBPad - C:\Documents and Settings\Richard Johnston\My Documents\Software\Basic\SmallBasic\Conditional.bas
File Edit Program About
Editor Help Output Text Output
The sum of the odd integers that are in the first 100 integers is 2500
* DONE *

Exercises

1. Write a program that calculates the Thevenin equivalent resistance R_i given $V_{Unloaded}$, V_{Loaded}, and R_L.
2. Write a program to add the sum of the even integers between 150 and 1000.
3. Write a program to calculate the roots of the quadratic equation

$$a\cdot x^2+b\cdot x+c=0$$

Note that the quadratic formula

$$r_{1,2}=\frac{b}{2\cdot a}\pm\frac{\sqrt{b^2-4\cdot a\cdot c}}{2\cdot a}$$

fails if a=0, and in that case the single root must be written

$$r=\frac{-c}{b}$$

The quadratic formula also fails if $b^2<4\cdot a\cdot c$ in that case the roots must be reported as

$$r_{1,2}=\frac{-b}{2\cdot a}\pm\frac{\sqrt{4\cdot a\cdot c-b^2}}{2\cdot a}\cdot i$$

where i is the complex operator $i=\sqrt{-1}$ (These are called complex roots.)
Notice that the quadratic formula also technically fails if $b^2=4\cdot a\cdot c$ since we then have only a single root (called a repeated root) given as

$$r = \frac{-b}{2 \cdot a}$$

Your program must take into account all four possible cases:

1. a=0 (single root)
2. $a \neq 0$ and $b^2 = 4 \cdot a \cdot c$ (repeated root)
3. $a \neq 0$ and $b^2 < 4 \cdot a \cdot c$ (complex roots)
4. $a \neq 0$ and $b^2 > 4 \cdot a \cdot c$ (real roots as calculated by the quadratic formula in the usual form)

Your program should report what kind of roots (real, complex, repeated, or single) are calculated as well as reporting the root(s) themselves.