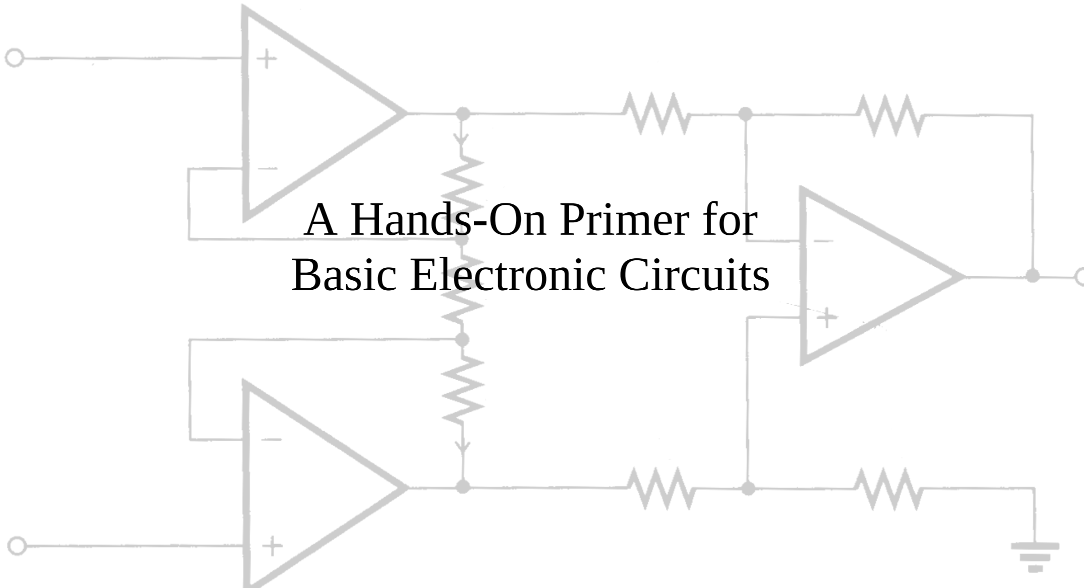
Principles of
ANALOG ELECTRONICS



A Hands-On Primer for
Basic Electronic Circuits

David B. Green
Natural Science Division
Pepperdine University

Table of Contents

SOME DEFINITIONS

Voltage: Voltage is the magnitude of the potential difference between a positively charged terminal (or plate) and negatively charged terminal (or plate). Voltage can be supplied either by a cell (often incorrectly called a "battery"), a battery (a group of cells in series or parallel), or power supply.

Battery schematic symbol

Unit of voltage is the volt (*V* or J/C) and the customary algebraic symbol for voltage is *E* (or *V*).

Current: The definition depends on who you ask. For this course, it is the direction of electron flow (from negative to positive) in an electric field. Often, current is described as the flow of electron "holes" from positive to negative.

Unit of current is the Ampere (*A* or C/s) and the customary algebraic symbol for current is *I*.

Resistance: Resistance is a property of a conductor, semiconductor, or insulator which opposes current flow. Often, the energy lost due to resistance manifests itself as the production of heat in the resistive device.

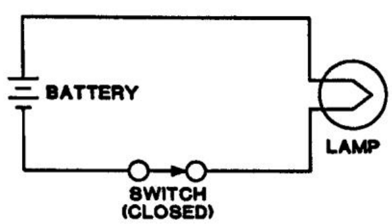
Resistor schematic symbol

Unit of resistance is the ohm (Ω or J s/C^2) and the customary algebraic symbol for resistance is *R*.

Power: Power is a catch-all term to describe the rate at which work is performed. Power is often dissipated by a circuit in the form of heat, although the production of light, sound, etc. can also result from the power output of a circuit.

Unit of power is the watt (*W* or J/s) and the customary algebraic symbol for power is *P*.

Load: Generally, a load is a device in the electric circuit which can perform useful work. For example, the light bulb in this circuit is the load.

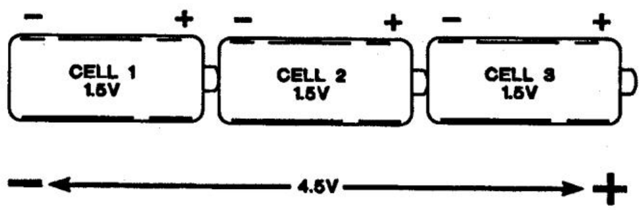


CONNECTING CELLS AND BATTERIES

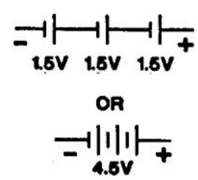
Cells can be connected together to increase the voltage or current rating. There are three different useful ways that cells or batteries can be connected. These are series, parallel, and series-parallel. We will examine each of these methods in some detail.

Series Connection

In the 12-volt automobile battery, six cells are connected together so that the individual cell voltages add together. In the 6-volt battery, three cells are connected in the same way. This arrangement is called a *series* connection and is shown in the figure below as it occurs in a three-cell flashlight. The cells are connected so that the positive terminal of the first connects to the negative terminal of the second; the positive terminal of the second connects to the negative terminal of the third; etc.



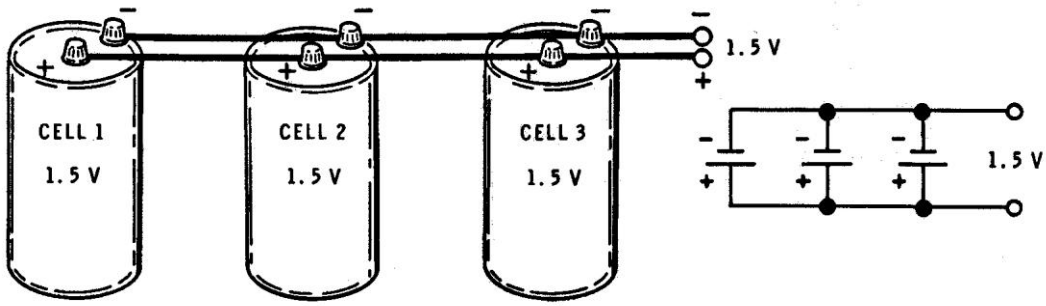
This is a series connection because the same current flows through all three cells. It is an aiding connection because the voltages add together. Since the individual voltage of each cell is 1.5 V, the overall voltage is 4.5 V. The schematic diagram for this connection is shown on the right.



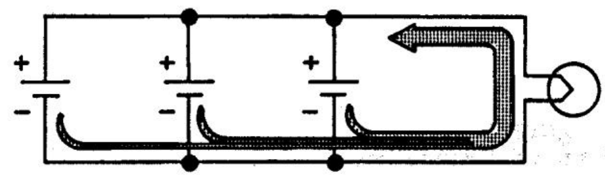
With the series connection, the total voltage across the battery is equal to the sum of the individual values of each cell. However, the current capacity of the battery does not increase. Since the total circuit current flows through each cell, the current capacity is the same as for one cell. If the cells have differing current capacities, the total available current is, for all practical purposes, the capacity of the lowest-capacity cell. So, for example, mixing 1.5 V “D-cells” with 1.5 V “AA-cells” in a series connection will increase the voltage of the battery to the sum of the cells but the total current capacity will be limited to that of the AA-cells.

Parallel Connection

We have seen that the series connection of cells increases the output voltage but not the current-supply capabilities of the cells. However, there is a way to connect cells so that their current capabilities add together. This is called a parallel connection and is shown in the next figure. Here, like terminals are connected. That is, all the positive terminals are connected together as are all the negative terminals.

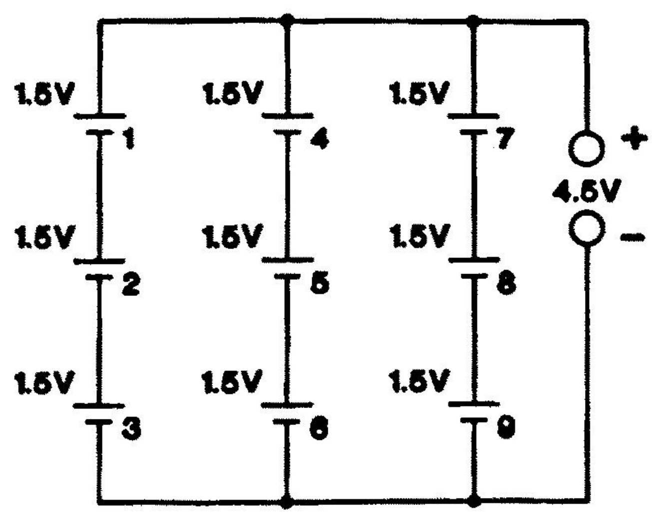


The following schematic diagram shows why the current capacities of the cells are added together. Notice that the total current through the lamp is the sum of the individual cell currents. Each cell provides only one third of the total current. Thus, the total current capacity is three times that of any one cell. However, connecting the cells in this way does not increase the voltage. That is, the total voltage is the same as that for any one cell. If 1.5 –volt cells are used, then the total voltage is 1.5 volts. Generally, it is unwise to place cells of differing voltages in parallel.



Series-Parallel Connection

When both a higher voltage and an increased current capacity are required, the cells are connected in series-parallel. Fore example, suppose we have four 1.5-volt cells and we wish to connect them so that the voltage is 3 volts and the current capacity is twice that of any one cell. We can achieve this by connecting the four cells as sown in the figure on the right. To achieve 3 volts, cells 1 and 2 are connected in series. However, this does not increase the current capacity. To double the current capacity we must connect a second series string (cells 3 and 4) in parallel with the first. The result is the series-parallel arrangement shown.



VOLTAGE DROP

Voltage Drop

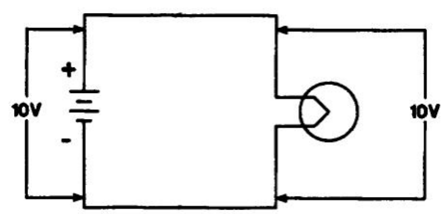
A cell, battery, or power supply produces a voltage. When a load is connected across the power source, electrons flow through the load. The fraction of the source voltage removed by the load is called a *voltage drop*. A voltage drop is expressed in volts.

Voltage Rise

In an electrical circuit a "voltage rise" is a voltage which is provided by a voltage source. A voltage rise is conventionally called a "source voltage".

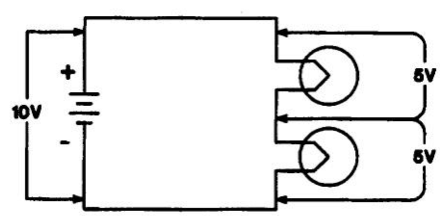
Voltage Drops Equal Voltage Rises (Kirchhoff's Voltage Law)

Consider a 10-volt battery with a light bulb connected across it.



The battery provides a source voltage of 10 volts. As electrons flow through the lamp, a voltage drop is developed across it. Since the lamp consumes the same amount of energy that the battery provides, the voltage drop across the lamp is equal to the voltage rise across the battery. That is, the voltage drop is 10 volts.

Now consider a circuit in which two electrically identical light bulbs are connected in series across a 10-volt battery.



Each bulb drops part of the 10 volts supplied. Since the two lams are identical, each will drop half of the supplied voltage as shown. If the two lamps are not identical, one bulb will drop more voltage that the other. In either event, this property can be summed up in **Kirchhoff's Voltage Law:**

The sum of the voltage drops will always equal the sum of the voltage sources.

CONCEPT OF GROUND

Among the most important principles in the study of electronics is the concept of *ground*. Originally, ground was just what the name implies – the Earth. Earth is arbitrarily assigned to have zero potential. Thus, ground (or earth) is the reference point to which voltages are most often compared. Many electrical appliances in your home are grounded. This is especially true of air conditioning units, electric clothes dryers, and washing machines. Often this is done by connecting a heavy wire directly to a cold water pipe which is buried deep in the Earth (ground). In other cases, a third prong on the power plug connects the metal frame to ground. The purpose of this is to protect the user in case a short circuit develops in the appliance. It also places the metal parts of different appliances at the same potential so that you are not shocked by a difference in potential between two appliances. This type of ground is sometimes called *earth ground*.

However, there is a slightly different type of ground used in electronics. For example, a certain point in a small transistor radio is called *ground* although the radio does not connect to the Earth in any way. This is the concept of ground with which we will be primarily concerned in this course. In this case, ground is simply a zero-volt reference point within an electric circuit. In most larger pieces of electronic equipment the zero reference point or ground point is the metal frame or chassis on which the various circuits are constructed. All voltages are measured with respect to this chassis.

In your automobile, the chassis or metal body of the automobile is considered ground. If you look closely at the straps leaving the battery you will see that one wire connects directly to the metal frame of the car. This point is considered to be ground as is every other point on the metal frame.

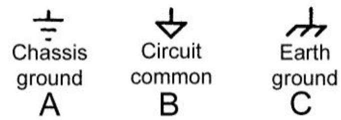
In electronics, ground is important because it allows us to have both negative and positive voltage. Up to now we have been concerned only with relative voltages between two points. For example, a 6-volt battery has a voltage between its two terminals of 6 volts. We do not think of this as +6 volts or -6 volts but, rather, simply 6 volts.

However, the concept of ground allows us to express negative and positive voltages. Remember *ground* is merely a reference point which is considered to be zero volts or neutral. If we assume that the positive terminal of a 6-volt battery is ground, then the negative terminal is 6 volts *more* negative. Thus, the voltage at this terminal with respect to ground is -6 volts.

On the other hand, if we assume that the negative terminal of the battery is ground, then the positive terminal with respect to ground is +6 volts. Notice that the battery can produce -6 volts or +6 volts depending on which terminal we call ground.

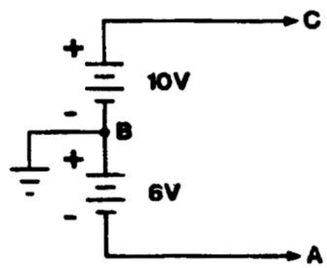
Many small electronic devices such as calculators, transistor radios, etc. do not have metal frames. Instead, all components are mounted on a printed circuit board. Here, ground is nothing more than an area of copper on the board. However, as before, all voltages are measured with respect to this point. In this case, ground is simply a common reference which is a specified starting point for measuring voltages.

The most widely used schematic symbols for ground are



Symbol A is used most commonly for “chassis ground” although it is often used synonymously with circuit “common” (symbol B). Symbol C is used to indicate that the circuit has been Earth grounded.

The presence of a ground in a circuit provides a method for constructing a “bipolar” power supply. Examine the following circuit.



Here, two batteries are connected in series, with the ground connection between them. Thus, the zero reference is at point B. Since the top battery has a voltage of 10 volts, the voltage at point C, with respect to ground, is +10 volts. The lower battery has a voltage of 6 volts. Because the positive terminal is connected to ground, the voltage at point A, with respect to ground, is -6 volts.

We sometimes loosely speak of the voltage at a particular point. Actually, voltage is always the measure of the potential difference between two points. Thus, in the figure above, when we speak of the voltage at point A, what we really mean is the voltage between point A and ground.

ELECTRICAL MEASUREMENTS I

Voltage

Introduction

The voltages used in these investigations are provided by the ±DC power supply on the Trainer. A variable positive voltage is provided between the positive terminal and the ground terminal. The value of this voltage can be adjusted by the positive (+) voltage control knob.

Similarly, a variable negative voltage is provided between the negative terminal and the ground terminal. This voltage can be adjusted by the negative (–) voltage control knob. Notice that the ground terminal is common to both supplies. Generally, voltages are measured with respect to this point.

Before you begin this investigation, be very sure you understand how to use your digital multimeter. Voltmeters vary greatly in their ranges, scales, and capabilities. Most meters can be damaged by incorrect use even at the low voltages used in the Trainer.

The voltmeter is typically a high impedance[1] device so as to not "load" a circuit under test. Voltages can be measured at the source (voltage rise or source voltage), across a electronic component (voltage drop), or at a point with the other lead at "ground".

Investigation 1: Measuring DC Voltage with the DMM

Materials Required

Trainer
Digital multimeter (DMM)

Procedure

1. Switch the Trainer on. Turn both of the voltage supply controls completely counterclockwise (lowest voltage settings).

2. Insert a piece of hook-up wire into the positive voltage source terminal and another into the ground terminal[2].

3. Set up the digital multimeter (DMM) to read DC (direct current) voltage with a maximum full-scale setting of at least 20 V. Some DMMs are "autoranging" and need only be set to read DC voltage.

---

[1] For now, impedance can be thought of as simply resistance.
[2] There are multiple ground terminals on the Trainer. All ground terminals are "common" so you may use the most convenient terminal.

4. Clip the black (negative) lead of the DMM to the ground wire and red (positive) lead to the positive voltage supply wire.

5. Record the minimum voltage attainable with the positive power supply. Remember to prefix this voltage with a plus (+) sign since it is more positive than ground.

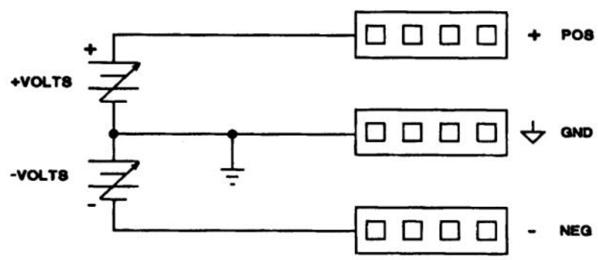
6. Turn the positive supply control about $^1/_3$ fully clockwise and measure the voltage. Notice how fine you can control the output voltage of the power supply.

7. Turn the positive supply control fully clockwise and measure the voltage. This is the maximum positive supply voltage the Trainer can produce. Record the maximum voltage attained.

8. Transfer the red lead to the negative power supply terminal and repeat steps 5-7. Notice that the DMM indicates that the voltage is negative with respect to ground.

Discussion

If you think of the power supplies on the Trainer as batteries, they are connected as shown here.



You should recognize this as two "variable batteries" in series with the ground assigned at the junction of the two supplies, as in the earlier description of the bipolar power supply.

Investigation 2: Measuring Voltage Drop with the DMM

Materials Required

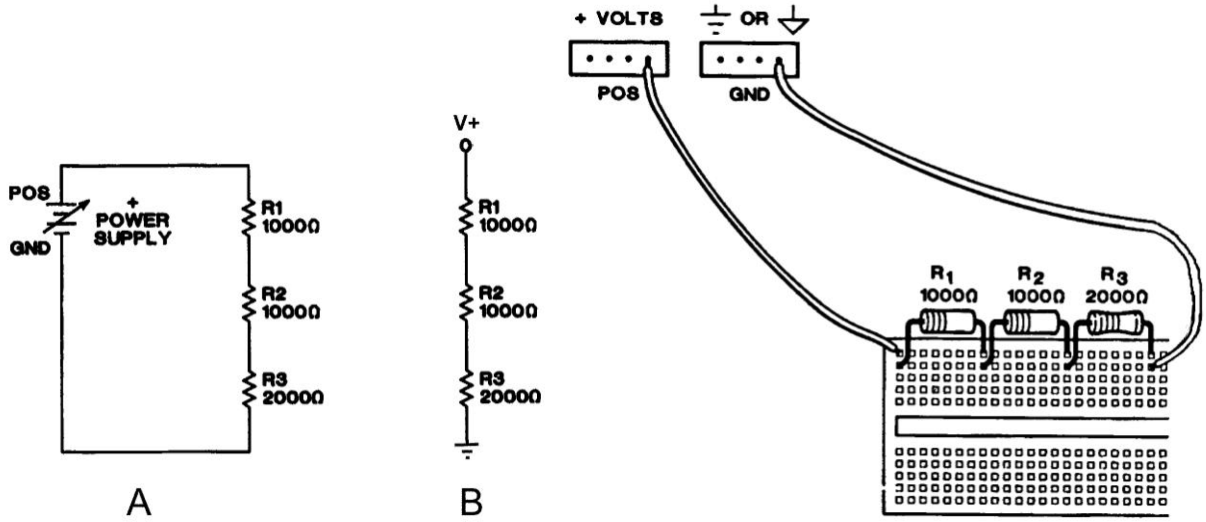
Trainer
Digital multimeter (DMM)
One 2000Ω resistor (red-black-red)
Two 1000Ω resistors (brown-black-red)

Procedure

1. Ensure that the Trainer is switched off[3]. Set the power supply controls to their minimum settings.

[3] The Trainer should always be turned off when connecting components and when building or modifying circuits.

2. Construct the circuit shown in the schematic. Schematic A is the circuit assuming the power supply is a "variable battery" while schematic B is the equivalent circuit using conventional schematic notation when a power supply is used. A view of a typical breadboard with installed components is shown on the right.



3. Set the DMM to read voltage with a maximum full-scale setting of at least 12 V. Clip the black lead to ground and the red lead to the positive supply terminal.

4. Turn on the Trainer and adjust the positive supply voltage to exactly 10 V.

5. Remove the voltmeter leads from the positive supply and ground. Attach the voltmeter leads across R1, observing correct polarity; that is, the positive lead should connect to R1 on a path electrical path that leads to the (+)-terminal.

6. Record the voltage drop across R1. Generally, voltage drops are magnitude only – a voltage drop is an unsigned potential.

7. Measure and record the voltage drop across R2 and R3 in a similar manner. Be sure to always observe correct polarity.

8. Sum the three voltages drops together. What is the sum of the voltage drops in this series resistance circuit? How would the voltage drops across each resistor be affected if they were rearranged in a different order? (You may wish to try this experiment.)

Leave the circuit assembled for Investigation 3.

Discussion

If the measurements were carefully taken, then you should have validated Kirchhoffs Voltage Law. These measurements are fundamental to virtually every circuit you build and provide a primary source of information for troubleshooting marginally functioning or nonfunctioning circuits.

You probably noticed that measuring the source voltage at the source terminals is no different that measuring the total voltage drop across the series network (*i.e.*, placing the test leads across R1 and R3). Another important aspect of voltage measurements is that, for all practical purposes, the voltage drop along a wire is zero so it does not matter where along the wire you place the test lead.

Investigation 3: The Effect of a Short or Open Circuit on Voltage Drop

Materials Required

Trainer
Digital multimeter (DMM)
One 2000Ω resistor (red-black-red)
Two 1000 Ω resistors (brown-black-red)

Procedure

1. Use the same source voltage and circuit assembled in Investigation 2.

2. While measuring the voltage drop across R3 (2000 Ω resistor), connect a short piece of hookup wire across R1 (1000 Ω resistor). This is called "short circuiting" (or just "shorting") the 2000 Ω resistor.

3. What was the effect of shorting R1 on the voltage drop across R3?

4. Remove the short from across R1. While measuring the voltage drop across R3, carefully pull out from the breadboard the resistor lead of R1 which connects to R2. This creates an "open circuit" at the R1-R2 junction.

5. What effect does this open circuit have on the voltage drop across R3?

Discussion

While this is a simple, perhaps trivial, investigation, it demonstrates two extremely important principles: 1) a short circuited component can cause increased voltage drops in other components, and 2) an open circuit can cause the circuit to quit functioning.

In the case of a short circuited component, the stress placed on the other components a circuit may exceed their operational tolerance and cause unexpected or unpredictable changes in their electrical characteristics or cause premature – sometimes startling – failure. In the case where the short results in the power supply's positive terminal being routed directly to ground or directly to the negative terminal (or *vice versa),* the results can be catastrophic to the power supply. In high voltage or high current power supplies, the results of a short may be life threatening. For example, uncounted house fires have been started and untold numbers of people killed from exposed 110V household electrical wires coming in contact with each other and creating sparks or heat.

Open circuits can be caused by numerous possibilities including component failure, broken conductors, components or wires poorly seated in the breadboard sockets, or components or wires in the wrong breadboard sockets, just to name a few. Open circuits in complex circuits or circuit assemblies designed to save space can be very frustrating to locate.

ELECTRICAL MEASUREMENTS II

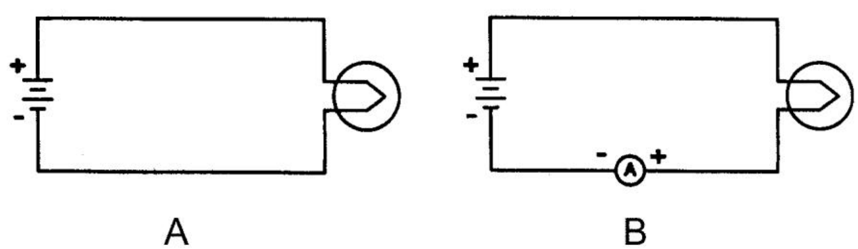
Current

Introduction

As we saw in the previous section, voltage is measured across a component or power source using a high input impedance (resistance) voltmeter. The voltmeter must have high input impedance to prevent the voltmeter from changing the load characteristics of the component under test.

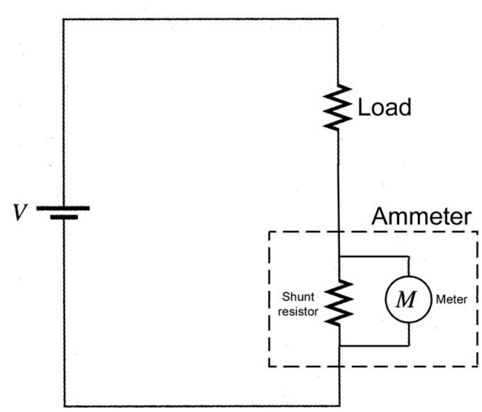
Current, on the other hand, is measured with the meter in the circuit. As should become clear shortly, the ammeter has a very low input impedance so as to not change the load characteristics of the circuit.

Consider circuit A in the following diagram. If we wish to know the current flowing through the circuit, we need to break the circuit and insert the ammeter in-line with the load (in this case, a lamp). Circuit B shows one possible placement of the ammeter. Notice that the ammeter, like the voltmeter, must be placed in-circuit while observing the correct polarity.



Ammeters are very delicate instruments and can be quickly destroyed if current in the circuit exceeds the current-carrying capability of the meter. For this reason, most modem digital multimeters have a fuse in their ammeter circuitry to protect the instrument from excessive current.

The ammeter itself is actually just an extremely sensitive voltmeter (calibrated in amps, milliamps, or microamps) measuring the voltage drop across a low resistance "shunt" resistor (1Ω or less). The basic ammeter circuit is shown in the dashed box in the diagram.



The shunt resistor, having a low resistance, changes the load in the circuit by a negligible amount thus appearing to the circuit as essentially a short. At the same, the low resistance of the shunt resistor makes the instrument susceptible to damage at high current. An ammeter must never be connected across power supply terminals or across any electronic component in a circuit – damage to the meter or a blown fuse **will** result.

Investigation 4: Measuring DC Current with the DMM

Materials Required

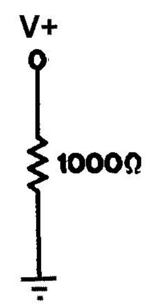
Trainer
Digital multimeter (DMM)
1000Ω resister (brown-black-red)

Procedure

Remember to record all of your data.

1. Ensure that the Trainer is switched off. Set the power supply controls to their minimum settings.

2. Construct the simple circuit shown in the following schematic.



3. Set the DMM to read voltage with a maximum full-scale setting of at least 12 V.

4. Turn on the Trainer and adjust the positive supply voltage to 5 V. **Turn off the Trainer.**

5. Remove the voltmeter leads from the supply and ground. Set the DMM to read current (mA) and adjust the sensitivity to the maximum milliampere range[4]. Break the circuit between the positive supply and the resister and, while observing polarity, attach the ammeter in-line from the (+)-terminal to the resistor.

6. Turn on the Trainer. If the current reading on the DMM is too low to give good accuracy, increase the sensitivity in steps until sufficient accuracy is obtained. Avoid an overload condition.

[4] When measuring current, it is important to always start at the highest range and increase the sensitivity in steps to avoid damaging the meter.

7. Increase the voltage until the current is exactly doubled. If necessary, change the DMM sensitivity to avoid an overload condition.

8. Turn off the Trainer. Reconnect the resistor to the (+)-termjnal.

9. Set the DMM to read voltage (remember to move the test leads to the appropriate jacks on the DMM if necessary). Turn on the Trainer and measure the new voltage.

How does the initial and final voltage compare? If the initial current was measured between the resistor and ground, how would it compare to the current measured between the (+)-terminal and the resistor?

10. Measure and record the current through the resister at 5 different voltages[5]. Create a graph of Current (mA) vs. Applied Voltage (V). What is the mathematical relationship between current and voltage?

Discussion

Voltage and current measurements provide much of the necessary fundamental measurements for performing a large number of tests on circuits. In this simple investigation, you should have observed that current and voltage are related (and directly proportional). In sections to follow, we will quantitatively explore the relationship of voltage and current.

[5] If you have two DMMs available, both measurements can be determined simultaneously; otherwise, follow the preceding procedure.

RESISTANCE

Resistance is a property of a conductor, semiconductor, or insulator which opposes current flow. Metallic conductors generally have low resistance; that is, they offer very little opposition to current flow; Insulators, such as plastic, rubber, and glass have very high resistances. Semiconductors, such as silicon or germanium, have intermediate resistances between conductors and insulators. The property of resistance (or, conversely, conductivity) can be explained at the atomic/molecular level and will be left to the reader to explore in greater depth.

The unit of resistance is the *ohm*, named for the German Physicist Georg Simon Ohm who discovered and described the relationship between voltage, current, and resistance. The ohm (Ω) is a complex unit ($J \cdot s/C^2$) which, for most people, has no intuitive meaning. However, most people can relate to the resistance of copper wire (which is exceedingly low). A copper wire 18 m long with a cross-sectional diameter of 0.65 mm is about 1 Ω.

Fixed-Value Resistors

A resistor is an electronic component which is designed to have a specified resistance. There is an exception to this 4efinition – the variable resistor – which will be discussed later. Four common types of resistors are manufactured: carbon-composition, carbon-film, metal-film, and wire-wound resistors. Each has its own advantages and disadvantages.

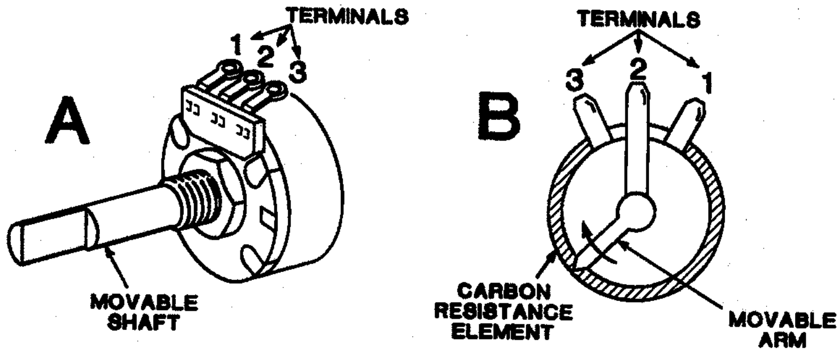
Carbon-composition resistors are manufactured using various ratios of carbon and an insulating binder hot-molded into a small thermal resistant polymer tube with metallic leads in intimate contact with the resistor material at each end of the tube. Composition resistors are inexpensive and suited to low current applications where the device does not have to dissipate large quantities of heat. Like most substances, carbon has a positive thermal coefficient of resistance; that is, the resistance increases with increasing temperature. Thus, the composition resistor should not be used in critical circuits where temperature variations will adversely affect circuit operation.

Metal-film and carbon-film (also called deposited-film) resistors are produced by depositing a metal or carbon film onto a nonconductive rod followed by scoring a helical groove from one end of the rod to the other. The wider the groove, the less the amount of conductive material is left on the nonconductive rod and the higher the resistance of the device. Metal film resistors suffer less from resistance changes due to temperature changes than composition resistors. Metal film resistors are also typically slightly more expensive than composition resistors but the cost difference is rather insignificant. The metal film resistor is an excellent general purpose device useful in all but the most demanding circuits.

Wire-wound resistors are usually made by winding nichrome wire onto a ceramic form then covered with a hard protective coating. Since they can dissipate large amounts of heat, wire-wound resistors can be used in high voltage/current applications without extreme changes in their resistance. Wire-wound resistors can also be produced with extremely precise resistances and are often used in high accuracy circuits such as those used in test equipment. Being coils of wire, circuit designers using wire-wound resistors have to be conscious of potential inductive effects of the resistor or purchase the more expensive non-inductive wire-wound resistor.

Variable Resistors

The most ubiquitous use of the variable resistor is the analog volume control on radios, stereo receivers/amplifiers, computer speakers, etc. An example of a variable resistor is shown in the following figure.



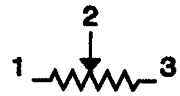
A. Example of a variable resistor. B. Internal connection of the carbon element to terminals 1 and 3 and the movable arm (wiper) on terminal 2.

The device shown is better and more properly known as a *potentiometer.* The resistance between terminals 1 and 3 is fixed by the resistive element but the resistance between terminals 1 and 2 is variable and depends upon where the wiper arm (movable arm) is in relation to terminal 1. The closer the wiper is to terminal 1, the lower the resistance between terminals 1 and 2. The same situation exists for terminal 2 and 3.

When used in the *rheostat* mode, only terminals 1 and 2 (or 2 and 3) are used. We will see later when to use a variable resistor in the potentiometer mode or rheostat mode.

The resistive element need not be made of carbon – wire-wound elements are used as well. Potentiometers of this type are known as – you guessed it – wire-wound potentiometers.

Potentiometers come in a very wide variety of sizes, shapes, and terminal geometries. Some can have a knob attached to a shaft, others require a screwdriver to turn the shaft. The one thing all potentiometers share in common is their schematic symbol:



The Resistor Color Code

Except for high-precision resistors, most resistors are imprinted with a color code to identify there value. The color code is easy to learn and, more importantly, with the first three color bands imprinted on the resistor its value can be discerned quickly. A fourth band indicates the tolerance in resistance value. The following table gives the entire color code for composite and film resistors.

Resistor Color Code				
	BAND			
	1	2	3	4[a]
Silver			$\times 0.01$	±10%
Gold			$\times 0.1$	±5%
Black	0	0	$\times 1$	
Brown	1	1	$\times 10$	
Red	2	2	$\times 100$	±2%
Orange	3	3	$\times 10^3$	
Yellow	4	4	$\times 10^4$	
Green	5	5	$\times 10^5$	
Blue	6	6	$\times 10^6$	
Violet	7	7	$\times 10^7$	
Gray	8	8		
White	9	9		
None				±20%

[a] Band 4 is the tolerance (maximum deviation from imprinted resistance)

Let's use the chart to determine the resistances for the resistors you have used in the previous investigations. One of the resistors had a color code of red-black-red.

The first band (red) is the first digit in the resistance (2).
The second band (black) is the second digit in the resistance (0).
The third band (red) is the resistance multiplier (×100).

Combining the numbers, we find the resistance to be

$$20 \times 100 = 2000\ \Omega \text{ or } 2\ \text{k}\Omega$$

The fourth band is the tolerance. Let's assume here that the fourth band is silver – a tolerance of ±10%. The resistor could have a value as low as 1800 Ω (2000 Ω - 200Ω) or as high as 2200 Ω (2000 Ω+200 Ω), which for most applications in this course is sufficiently accurate.

The other resistor used in previous investigations had a color code of brown-black-red. The code translates to

Brown = 1
Black = 0
Red = ×100

Thus, the value of this resistor is 1000 Ω (=1 kΩ).

ELECTRICAL MEASUREMENTS III

Resistance

Introduction

The instrument used for measuring resistance is called an ohmmeter. All DMMs have an ohmmeter in addition to their voltmeters and ammeters. The ohmmeter function on the DMM uses an internal battery to apply a small voltage to the device under test then measures the current which passes through the device. The current measured is calibrated and displayed in ohms.

The ohmmeter must never be connected to a circuit or component to which power is applied. The ohmmeter being an ammeter in function can be easily burned out.

Investigation 5: Measuring Resistance with the DMM

Materials Required

Trainer
Digital multimeter (DMM)
5 assorted resistors
Two 1000 Ω resistors (actually, any two resistors from 500-2000 Ω will work fine)
10 kΩ - 100 kΩ potentiometer (you can use the potentiometer on the Trainer)

Procedure

Record all of your data.

1. Set up the DMM to read resistance.

2. Examine a resistor's color and determine its resistance and tolerance. Set the DMM resistance range to a sensitivity which will include the resistor’s value.

3. Attach the ohmmeter leads to the resistor. Polarity in this case is unimportant. Read the resistance off the meter display.

4. Repeat for all of the resistors.

Do the resistances measured fall within the stated tolerances of the resistors? Do any exceed the tolerance?

An interesting experiment is to obtain 20-30 resisters of the same value and measure their individual resistances, then construct a histogram of the resistances.

5. Attach the leads to terminals 1 and 3 of the potentiometer. Record the resistance. This is the resistance of the resistive element of the potentiometer.

6. Move one of the leads to terminal 2 of the potentiometer. Adjust the potentiometer in both directions.

What is the maximum and minimum resistances obtainable?

7. Connect the two 1000 Ω resistors in series on the breadboard shown in schematic form:

8. Using the ohmmeter, determine the total resistance of the series circuit.

How does the total resistance relate to the individual resistances?

Discussion

The resistor color code makes it possible to quickly determine the value of a resistor without having to actually measure the resistance. With relatively rare exception, the measured resistance of the resistor will fall within the stamped tolerance.

Variable resistors will be used extensively during the course to control voltages and current flow.

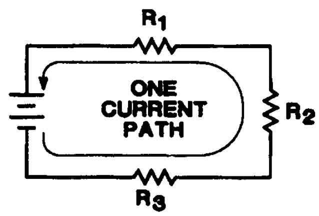
Finally, it should have been observed that, within stamped resistance tolerance, the resistance of a series network of resistors is the sum of the individual resistances, a concept to be covered in more depth in the next section.

RESISTOR NETWORKS

Resistors are often connected in series, in parallel, and in series-parallel combinations. In order to analyze and understand electronic circuits, we must be able to compute the total resistance of resistor networks.

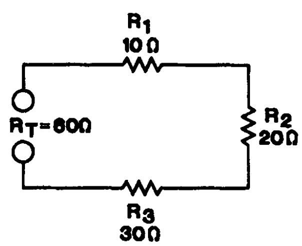
Resistors In Series

A series circuit is one in which the components are connected end to end, as shown in the following circuit.



It can be easily shown that the same current flows through all components. The current in the circuit must flow through all three resistors one after the other. Therefore, the total opposition to current flow is the total of the three resistances.

An example of three resistors connected in series is shown in the next figure.



The total resistance is called R_T and is the resistance of the entire series circuit. R_T is found by adding the individual resistor values together. That is:

$$R_T = R_1 + R_2 + R_3$$

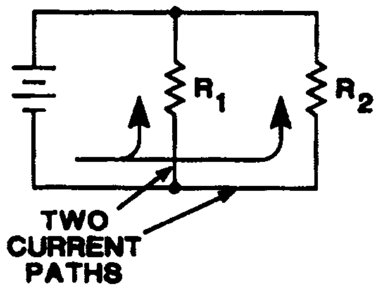
$$R_T = 10\Omega + 20\ \Omega + 30\ \Omega = 60\ \Omega$$

The three resistors in series have the same opposition to current flow as a single 60-ohm resistor.

This example used three resistors. The same principle holds true for any number of series resistors.

Resistors In Parallel

In parallel resistor network, components are connected across each other so that there are two or more paths for current flow.



Regardless of the size of R_1 and R_2, the total current provided by the current source will always increase when R_2 is placed in parallel with R_1 because a second current path is created. Obviously then, the total opposition to current flow decreases since more current now flows. Thus, when one resistor is placed in parallel with another, the total resistance decreases.

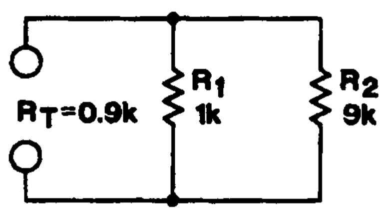
Without mathematical proof, there is a simple formula to determine the resistance of a parallel resistor network:

$$R_T = \left(\frac{1}{R_1} + \frac{1}{R_2} + ... + \frac{1}{R_n} \right)^{-1}$$

The above formula is normally used when more than two resistors are in parallel. However, the formula can be simplified if only two resistors are in the parallel network:

$$R_T = \frac{R_1 \times R_2}{R_1 + R_2}$$

Consider the following parallel resistor circuit:



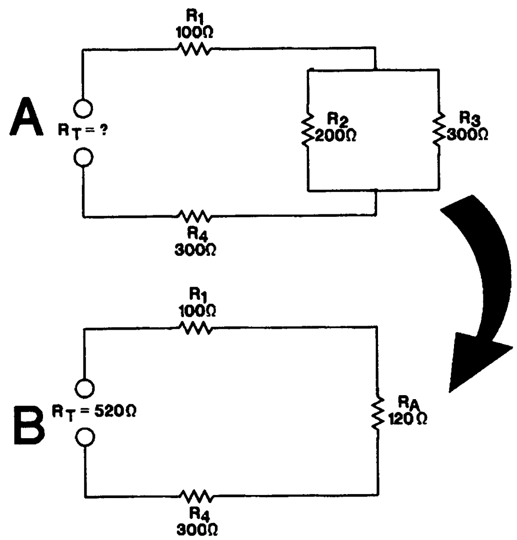
Using one of the above formulas, an equivalent resistance for R_1 and R_2 can be found.

$$R_T = \frac{1000\Omega \times 9000\Omega}{1000\Omega + 9000\Omega} = 900\Omega$$

Notice that the resistance of the parallel network is less than the smallest resistor in the network, which will always be the case. This observation helps to quickly verify that the calculation was performed correctly. If the parallel resistance calculates higher than the smallest resistance the technician should carefully repeat the calculation.

Series-Parallel Connections

In many circuits a parallel circuit is connected in series with one or more resistors as shown in circuit A of this schematic diagram.

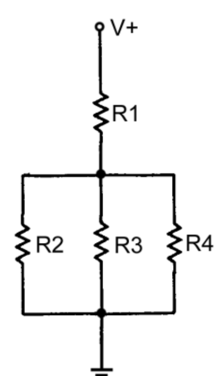


Even so, the total resistance is easy to calculate using the formulas shown earlier. The procedure is to compute an equivalent resistance for the parallel circuit first. Then this equivalent resistance is added to the series resistance values to give a total resistance of 520 Ω.

Investigation 6: Voltage Drops in a Series-Parallel Resistor Network

Materials Required

Trainer
Digital multimeter (DMM)
Two 1 kΩ resistors
Two 10 kΩ resistors



Procedure

Record all of your data.

1. Assemble on the breadboard the circuit shown at the right. $R_1 = R_4 = 1$ kΩ; $R_2 = R_3 = 10$ kΩ; V+ = 10V.

2. Measure and record the voltage drop across each resistor.

3. Measure and record the current at the following points:

V+ to R1
R1 to top of network junction of R2-R3-R4
Bottom of network junction R2-R3-R4 to ground
Through R2
Through R3
Through R4

4. When measurements are completed, leave the circuit assembled for later study.

What is the relationship of the sum of the currents through the parallel network to the current from V+ to R1?

Discussion

This investigation demonstrates several important electrical characteristics of resistor circuits. First, the voltage drop of one resistor in a parallel network is the same as all others in the parallel network. Second, current entering a network is equal to the current exiting the network. Third, the sum of the currents in a parallel network is equal to the to current entering (or exiting) the parallel network.

The first observation is merely another demonstration of Kirchhoff's voltage law. The second and third observations demonstrate **Kirchhoff's Current Law:**

The sum of the currents into a point in a circuit equals the sum of the currents out

OHM'S LAWS

Ohm's laws define the way in which current, voltage, resistance, and power dissipation are related. We have examined this qualitatively in prior sections but now let's examine the relationship quantitatively.

Investigation 7: Relationship of Voltage Drop and Current

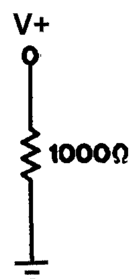
Materials Required

Trainer
Digital multimeter (DMM)
1 kΩ resistor

Procedure

Remember to record all of your data.

1. Construct the circuit shown in the following schematic. Set the supply voltage to 12 V.



2. Measure the current through the resistor.

3. Change the voltage either up or down and measure the voltage and current again. Perform the measurements for a total of 5 different voltages. If you have two DMMs you can assign one as a voltmeter and the other as an ammeter and acquire the data more quickly.

Make a plot of I vs. E and determine the slope of the straight line.

Discussion

Your first observation should be no more than review: as the voltage drop across the resistor increases, the current through the resistor increases. More significantly, the slope of the line derived from the E-I plot should be the measured resistance of the resistor.

The relationship of E, I and R is simply

$$E = I \times R$$

This humble relationship is known as Ohm's Law and will be used to help in designing every circuit from very simple resistor networks all the way to the most complicated circuits.

Let's explore the relationship some more. Examine the data acquired in Investigation 6. Using Ohm's law, calculate the current flowing through R1 based on the resistance of R1 and the voltage drop across R1. Is the calculated current similar (or identical) to the measured current? Similarly, use the voltage drop across R2, R3, and R4 to calculate the current through each of these resistors. Are the calculated currents similar to those measured?

POWER

In addition to the three basic electrical quantities (current, voltage, and resistance) a fourth quantity is also very important – **power.** Power is defined as the rate at which work is done. In other words, power refers to the amount of work done in a specific length of time. The production of power in a circuit is called dissipation.

Power is the work performed over time and possesses the units of J/s or watts. It is named in honor of James Watt who pioneered the development of the steam engine. When we speak of the power used by a circuit, it is in terms of the number of watts dissipated by the circuit.

Voltage, Current, Resistance; and Power

Power is directly proportional to both voltage and current. Voltage is the energy per coulomb (J/C) of charge accelerating an electron down a conductor and current is the number of coulombs of charge passing a given point per second (C/s). Thus, the formula for power is:

$$P = E \times I$$

Rearranging and substituting Ohm’s law into the equation for power yields the relationship of power with each of the electrical quantities:

$$P = I^2 \times R$$

and

$$P = \frac{E^2}{R}$$

Let's explore each of these relationships in a little depth.

Power Dissipation in Resistors

Resistors and most other electronic components dissipate power in the form of heat. In some cases, the heat produced is a desired result. For example, the purpose of the resistive element in a toaster, heating pad, or electric stove is to produce heat. However, in most electronic devices, the heat produced by resistors represents wasted power.

The power lost by the heating of resistors must be supplied by the power source. Since electrical power costs money, attempts are generally made to keep the power lost in resistors at a minimum.

Example 1:
Consider the case where a resistive load has a voltage drop of 110 V and 10 A of current are passing through the load. What is the power dissipation for this circuit?

$$P = E \times I = 110\text{ V} \times 10\text{ A} = 1100\text{ W}$$

Example 2:
The maximum power rating of a particular 500 Ω resistor is ½-watt. In a circuit, it produces a voltage drop of 10 V. Is the maximum power dissipation rating exceeded in this circuit?

$$P = \frac{E^2}{R} = \frac{(10\text{V})^2}{500\Omega} = 0.2\text{ W}$$

Since the device can dissipate 0.5 W safely, the 0.2 W being dissipated by the device does not exceed the operational parameters of the resistor.

Calculate the power dissipated by the Rl in Investigation 6 based on the current through R1 and its resistance.

Optional Investigation: Exploration of the Power Law by Calorimetry

Materials Required

DC power supply capable of producing 1 A current.
Digital multimeter (DMM)
100 Ω (1 watt) resistor
Soldering iron
Two polystyrene coffee cups with insulated lid
Thermometer
Stopwatch or timer

Procedure

Remember to record all of your data.

1. Solder ~20 cm lengths of hookup wire onto the leads of 1 W 100Ω resistor. Accurately measure the resistor.

2. Attach the resister to the power supply in such a way as to be able to measure current. Set the supply voltage to about 10 V. Accurately set the current to 0.100 A by fine adjustment of the voltage. Note the voltage, current, and resistance. Turn off the power supply.

3. Assemble a colorimeter by nesting two polystyrene coffee cups together. Transfer 25 mL of deionized water into the calorimeter and cover with an insulated lid. Dip the 100Ω resistor into the water[6].

[6] The use of deionized water, being non-conductive, will minimize the need to insulate the exposed wires and the leads of the resistor.

4. Pass the thermometer by the lid and measure the temperature of the water.

5. While gently swirling the water in the calorimeter turn on the power supply and verify that the current through the resister is still 0.100 A. Adjust as necessary. The power dissipated by the resistor will heat the water.

6. Allow the circuit to operate for at least 5 minutes, swirling frequently.

Recall that the amount of heat energy absorbed by a substance can be calculated by

$$q = mc\Delta T$$

where q is the amount of heat in joules absorbed by the substance, m is the mass of the substance, c is the specific heat capacity of the substance, and ΔT is the change in temperature of the substance. The total amount of heat energy electrically produced by the resistor is

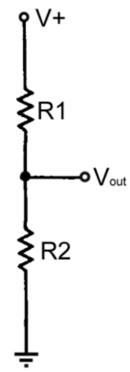
$$q = P \times t$$

where P is the power dissipated by the resistor and t is the amount of time the power was produced.

7. From the temporal interval for which the resistor was dissipating power, calculate the predicted temperature change for the water in the calorimeter. Compare this value to the experimentally determined temperature change.

A PRACTICAL RESISTOR CIRCUIT - THE VOLTAGE DIVIDER

Among the most used resistor networks is the *voltage divider*. The voltage divider is simply two or more resistors in series connected, often at the voltage source at one end of the resistor network and ground or the opposing source (in a circuit using bipolar power supply) at the other end. This circuit shows a basic voltage divider network. The voltage measured at V_{out} is taken as the voltage drop of R2 relative to ground.



Using Ohm's law, prove that the voltage measured at V_{out} is given by

$$V_{out} = V_{+} \times \left(\frac{R_2}{R_1 + R_2} \right)$$

This circuit and similar ones like it will be encountered frequently throughout this course.

Investigation 8: The Voltage Divider

Materials Required

Trainer
Digital multimeter (DMM)
4.7 kΩ resistor
10 kΩ resistor

Procedure

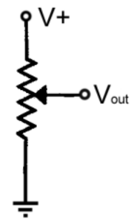
Remember to record all of your data. You may recall that you built a similar circuit in Investigation 2.

1. Construct a voltage divider as shown in the introduction to this section with R_1 = 10 kΩ and R_2 = 4.7 kΩ. Set the supply voltage to +12 V.

2. Measure the voltage drop across R_2.

Calculate the expected voltage drop across R_2. Is the calculated voltage produced by the voltage divider within the expected range, based on the tolerance of the resistors in the network?

3. Set up the following circuit using the l00 kΩ potentiometer provided on the Trainer with terminal 3 at V+ and terminal 1 at ground:



4. Set the source voltage to 12 V and measure V_{out} at terminal 2 of the potentiometer versus ground.

5. While measuring V_{out}, adjust the potentiometer first fully clockwise then fully counterclockwise. Observe the effect of adjusting the potentiometer on the voltage produced by this "adjustable" voltage divider circuit.

6. Adjust the potentiometer until the output voltage is exactly 5 V. Without changing the potentiometer setting, turn off the Trainer and remove the hookup wire from the V+ terminal and the ground terminal.

7. Using the DMM in the ohmmeter mode, measure and record the resistance of the potentiometer first between terminals 1 and 2, then 2 and 3. Also, measure the total resistance of the potentiometer between terminals 1 and 3.

Based on source voltage and the total resistance of the potentiometer, calculate the expected value of the resistance of the potentiometer between terminals 2 and the terminal attached to ground. Are the measured and calculated values of the resistance between these terminals similar?

Reattach terminal 3 to V+. Attach terminal 1 to V-. Set both power controls to their minimum setting and the 10 kΩ potentiometer about midway.

Turn on the Trainer and set V+ to 5 V and V- to –5 V.

While observing the voltage at V_{out} (relative to ground) turn the potentiometer control in both directions. Notice that you can achieve about +5 V at one extreme of the potentiometer and about -5 V at the other. Also you should have noticed that the voltage passes through zero at about the middle of the potentiometer travel.

Discussion

The voltage divider is a profoundly useful circuit and is found in virtually every useful electronic circuit in one form or another. You might recognize the "variable" voltage divider as demonstrated here being similar in function to the voltage control built in to the Trainer or the analog volume control on a radio, CD player, cassette tape player, etc.

The ability to “trim” the output voltage to + or - values will find use later when building amplifiers which require input "bias" adjustments.

ALTERNATING CURRENT

Introduction

So far, we have restricted our studies to direct current (DC) systems only. While important, DC circuits comprise only a small number of the possible basic useful circuits. It is essential that anyone studying the field of electronics have a thorough introduction to the principles of alternating current (AC). Far more electronic systems take advantage of the flexibility afforded by AC electronics; in fact, it is virtually impossible to avoid using it or, at least, having a circuit being affected (positively or adversely) by it.

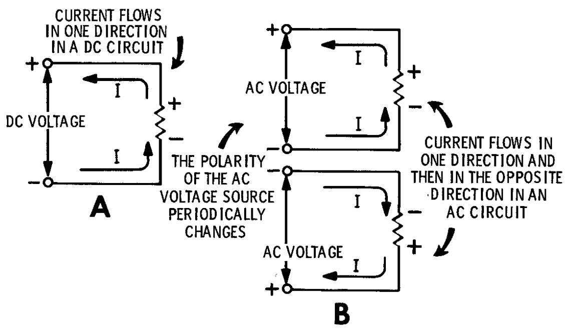


Figure A reviews how direct current flows through a load while Figure B shows how alternating current flows through a load at two points along a complete cycle. In one-half of an AC cycle, the current flows counterclockwise through the load. When the cycle reverses, the signs of the applied voltages invert and the current flows in the alternate direction.

Fortunately, Ohm's laws and the power laws remain unchanged for AC circuits. However, new ways of describing voltage, current, and power will need to be introduced, as we will see later in this section.

Why use AC at all? One example is in the transmission of power over a long distance. AC current can be transmitted along a wire over very long distances with far less loss than DC. AC voltages (or currents) can easily be converted to higher or lower voltages (or currents) using a device known as a *transformer*. To be certain, DC voltages can be stepped up (or down) as well but the process is far more complex and much less efficient. In fact, the usual method of stepping up DC voltages is to first electronically convert it to AC, step the voltage up with a transformer, then convert the AC back to DC with an increased voltage.

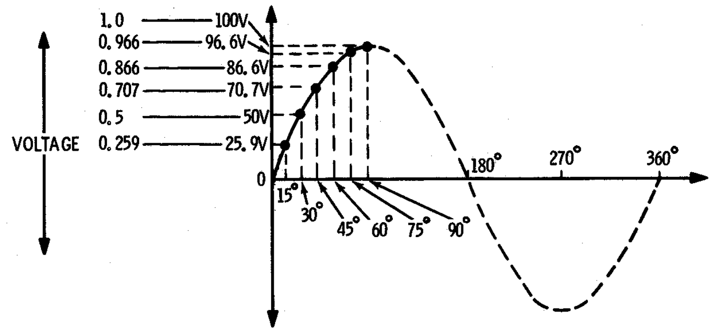
Alternating current can be easily converted to DC with nothing but a passive circuit (a circuit which does not require an external power source) but the opposite conversion requires an active circuit called an *oscillator*.

Alternating current also possesses a frequency of oscillation. As the frequency increases into the kHz and MHz range, the energy can be converted to electromagnetic energy, in the form of radiowaves, to carry information from one point to another without wires.

A common machine which produces AC is called a *generator*. It is left to the reader to study the construction and function of the AC generator.

The AC Cycle

The diagram shows one complete alternation of a sinusoidally alternating voltage[7]. During one-half of the cycle, the voltage applied to the load is positive while during the second half of the cycle, the voltage polarity is reversed. The cycle then repeats itself. At 0°, 180°, and 360° along the alternation the voltage applied to the load is momentarily zero.



AC Values

Since the voltage of an AC signal is constantly changing, there exists multiple ways of describing the AC voltage. In other words, you cannot simply say that the voltage of the sine wave above is 100 V without in some way specifying that this is the maximum value of the voltage.

Peak Voltage

In the diagram, the maximum voltage (peak voltage) during any cycle is 100 V. Thus, the peak voltage, V_p, of this signal is 100 V_p.

Peak-to-Peak Voltage

Often it is necessary to know the total height of the sine wave. The overall voltage of the sine wave from peak to trough is the peak-to-peak voltage, $V_{p\text{-}p}$. In the example, the peak-to-peak voltage of the sine wave is 200 $V_{p\text{-}p}$.

Average Voltage

Infrequently, the average voltage of a sine wave is of interest. The average peak voltage, $V_{average}$, is $0.636 \times V_p$.

RMS Voltage

When DC of a set voltage flows through a resistor, a certain amount of power is dissipated in the form of heat. Likewise, when AC with the same peak voltage flows through the same resistor, power is dissipated as heat but the amount of heat produced will necessarily be less. The result is that a higher AC peak voltage is needed to be equivalent to a specified DC voltage. Without proof, the effective voltage (called the root-mean-squared voltage or V_{rms}) is $0.707 \times V_p$. Thus, for the example above, the effective voltage of the sine wave is V_{rms} = 70.7 V. Root-mean-squared

[7] As we will see later, the alternations do not necessarily have to be sinusoidal.

voltage is used extensively when working with AC circuits especially when power dissipation is a concern.

AC Frequency and Period

The frequency, *f*, of the AC sine wave is defined as the number of complete cycles that occur per second. Frequency has the units of Hz, /s, s^{-1}, and cycles/second, all of which are just different ways to say the same thing. An AC source which produces 60 complete cycles per second (such as household current in the United States) is said to have a frequency of 60 Hz. An AM radio station may transmit on 640 kHz which means the transmitter electronics use an AC signal which alternates 640,000 times per second.

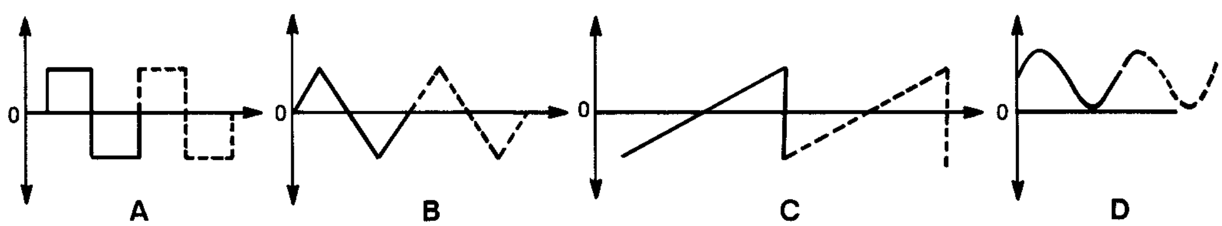
Period, T, is simply the time required to complete one AC cycle and is calculated as

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

and has the units of time.

Nonsinusoidal Waveforms

The sine wave is the most basic and widely encountered waveform. However, it is certainly not the only type of waveform that is used in electronics. The following examples show just a few of common nonsinusoidal waveforms frequently encountered in different types of oscillating circuits.



Nonsinusoidal waveforms

Square Wave

The square shape of the oscillating voltage (figure A) gives this waveform its name. It is really just positive to negative fluctuating DC. Changing the width of the positive or negative (or both) half of the cycle gives rise to a variety of different wave characteristics. The square wave is often used in timing circuits where fluctuation DC is more practical that sinusoidal AC.

Triangular Wave

As in the square wave, the triangular wave (figure B) gets its name from its shape. Triangular waves are usually used as electronic signals and control signals. Rarely is the triangular waveform used to transmit or provide electrical power.

Sawtooth Wave

The sawtooth waveform (figure C) is similar to the triangular waveform except that either the cycle ends with or starts with an abrupt voltage change. In the example shown, each cycle starts at a negative value and ends at a positive value. At the end of the cycle, the voltage returns discontinuously back to the starting value.

Wave Riding on DC

Very often, especially where certain types of amplifiers are concerned, an AC signal will be superimposed on a steady DC signal (figure D). In this case, the AC signal is said to be "riding" on the DC signal. The result is that the inflection points of the AC signal are not at zero volts.

ELECTRICAL MEASUREMENTS IV

AC Voltage

Introduction

Nearly all DMMs which can measure DC voltage can also measure AC voltage. Some are equipped to read peak voltage, peak-to-peak voltage, or rms voltage. It is critically important to know what voltage your meter is reporting.

While no damage will normally occur, it is important to remember to set the DMM to read AC voltage rather than DC voltage when AC circuits are being analyzed. Also, some AC voltmeters will only work for frequencies around 60 Hz. If unsure about operation or frequency response, consult the instruction manual for the meter.

Investigation 9: Measuring AC Voltage with the DMM

Materials Required

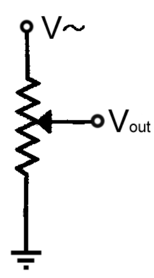
Trainer
Digital multimeter (DMM)
100 kΩ potentiometer (use the potentiometer on the Trainer)

Procedure

Record all of your data.

1. Set up the DMM to AC voltage.

2. Assemble a variable voltage divider circuit similar to the one in Investigation 8:



3. Use the 15 V AC terminal for the supply voltage.

4. Measure the AC voltage with respect to ground. Adjust the potentiometer to show that you can control the AC voltage in a similar manner as DC voltage.

ELECTRICAL MEASUREMENTS V

The Oscilloscope

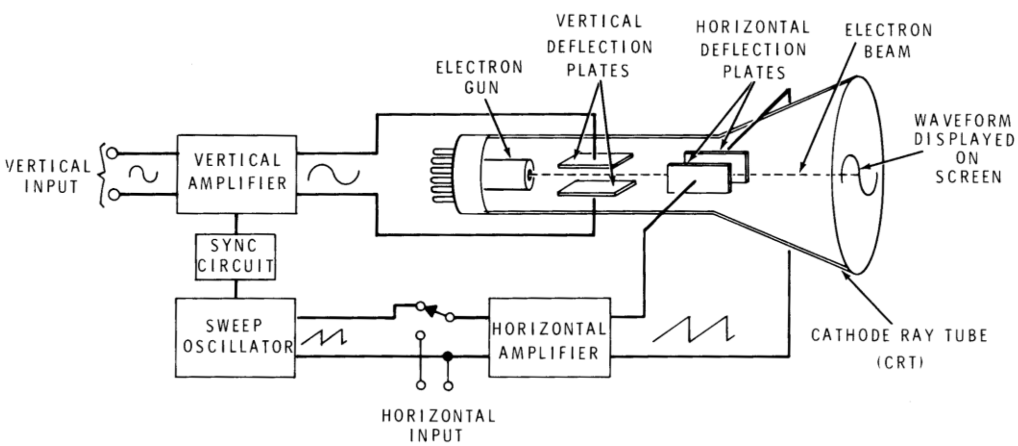
AC meters can provide a reasonably accurate measurement of current and voltage, but these instruments do not allow you to see what the AC signal actually looks like. They are usually calibrated to indicate the effective or rms value of a sine wave and, if they are used to measure nonsinusoidal waveforms, their scales do not provide true rms readings. Some AC meters have a frequency response which makes them inaccurate at frequencies higher than about 100 Hz.

When troubleshooting or analyzing an electronic circuit, it is often necessary to know exactly what an AC waveform looks like. In many cases it is necessary to know its peak or peak-to-peak voltage, its instantaneous values, and also its frequency or period as well as phase relationship. These measurements, among others, can be performed by using a instrument known as an *oscilloscope.*

The oscilloscope may be used to analyze any type of AC waveform and measure its most important electrical characteristics. The oscilloscope is one of the most important test instruments for measuring AC quantities. Although used primarily in AC circuit analysis, the oscilloscope may also be used to measure DC quantities as well.

Oscilloscope Operation

An oscilloscope is capable of measuring an AC or DC voltage and displaying the voltage graphically. The AC or DC voltage appears as a picture on a screen which is similar to the type of screen used in a television set. The modern oscilloscope contains a number of controls which are used to adjust the size and the number of complete waveforms that are displayed as well as where in the waveform the AC voltage is sampled, how the instrument is triggered to display the waveform, and other functions. Most oscilloscopes are calibrated so that the waveform presented on the screen can be visually analyzed and its most important characteristics can be determined. Some oscilloscopes can sample and electronically store a fast or transient waveform.



Block Diagram of an Oscilloscope

Examine the simplified block diagram of an oscilloscope. The oscilloscope has two input terminals which are used to measure an AC or DC voltage one of which is normally ground. The input resistance (impedance) of an oscilloscope is normally very high – 10 MΩ is not unusual. These terminals must be connected across the voltage source that is to be measured and they are generally referred to as the vertical input terminals. The AC voltage at these terminals is applied to an amplifier circuit which increases the amplitude or magnitude of the voltage before it is applied to a device known as a cathode ray tube or CRT. The CRT is the device which graphically displays the AC waveform being measured.

The CRT contains an electron gun and two sets of deflection plates. These components are mounted inside of a large evacuated glass tube which fans out at one end to form a screen which closely resembles the screen on a television picture tube. The electron gun produces a stream of electrons which are focused into a narrow beam and aimed at the CRT screen. When the beam strikes the screen, it illuminates a phosphorus coating on the screen so that a spot of light is produced. This electron beam must also flow between the two sets of deflection plates.

The AC voltage from the vertical amplifier is applied across the vertical deflection plates. This alternating voltage causes the plates to become positively and negatively charged and the polarity of these charges is continually reversed. The electrons in the beam are negatively charged and tend to deflect toward the positive plate and away from the negative plate, thus causing the electron beam to bend. Since the charges on the vertical plates continually change direction, the electron beam is deflected up and down thus causing a vertical trace to appear on the CRT screen. The height of this vertical trace will depend on the amplitude of the AC voltage being measured and the amount of amplification provided by the amplifier circuit.

If the electron beam was simply moved up and down, only a vertical line or trace would appear on the screen. Such a display could indicate the peak-to-peak amplitude of a waveform but still would not indicate the exact shape of the waveform. In order to show how the waveform varies, it is necessary to move the electron beam horizontally across the screen. This is accomplished by a circuit known as the sweep oscillator. This circuit generates an AC sawtooth waveform which is then amplified by a horizontal amplifier and then applied to the horizontal deflection plates. The sawtooth voltage increases at a linear rate from a negative peak value to a positive peak value and then almost instantly changes back to a negative value again. The positive and negative charges on the horizontal deflection plates vary in the same manner thus causing the electron beam to move from left to right across the screen at a linear rate and then immediately jump back to the left side and start over again.

If only the sawtooth waveform was applied to the horizontal plates with no voltage on the vertical plates, then only a horizontal trace would appear on the screen. Such a trace is thought of as the *time base* upon which the vertical signal can be made to ride. The beam simply moves from left to right in a specific period of time and then repeats this action again and again.

When the vertical AC voltage and the horizontal sawtooth voltage are both applied to the CRT, an AC waveform can be produced. As the beam moves from left to right at a linear rate with respect to time, the vertical AC voltage causes the beam to move up and down in accordance with the variations in AC voltage. If the time required for the beam to move across the screen from left to right is equal to the time required to generate one cycle of the AC input voltage; one cycle of the AC waveform will appear on the screen.

To insure that the input AC waveform and the sawtooth waveform are properly synchronized, a synchronization or "sync" circuit is included in the oscilloscope circuit. This circuit, known as the *trigger,* samples the incoming AC signal and produces a control signal which is applied to the sawtooth oscillator so that the sawtooth begins its cycle at the proper time.

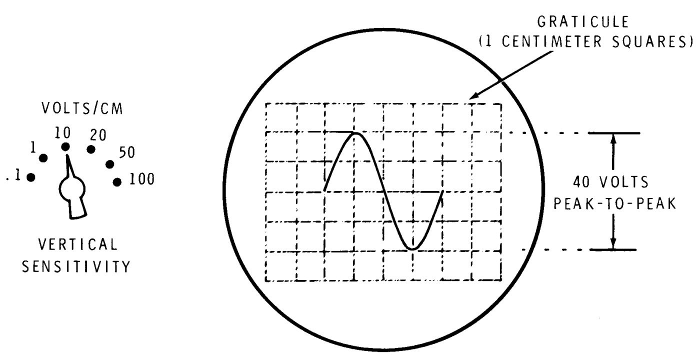
The input AC waveform and sawtooth waveform do not simply occur just once. These waveforms must occur repeatedly in order to produce a picture on the screen. In other words, the electron beam follows the pattern of the waveform again and again and this results in a constant picture on the screen. The phosphor on the screen produces light for only an instant after the electron beam strikes it and moves on. Therefore, this constant repetition is necessary to produce a pattern that is constantly illuminated.

Using the Oscilloscope

The oscilloscope can be used to observe various types of AC waveforms and it can be used to measure important AC values.

Measuring Voltage

Since the oscilloscope displays an entire AC waveform, it can be used to determine instantaneous values as well as peak and peak-to-peak values. A typical oscilloscope display is shown below. Notice that the screen of the oscilloscope is marked with vertical and horizontal lines which form squares. This grid pattern is commonly referred to as a *graticule.* The squares are usually 1 centimeter high and 1 centimeter wide and are used in much the same way as a sheet of graph paper.

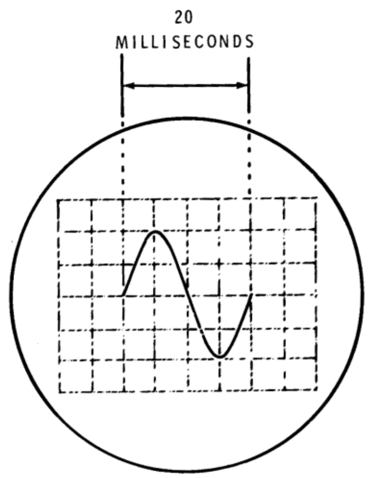


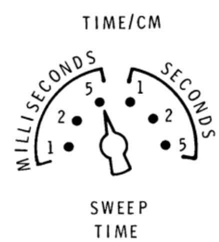
Observe the AC waveform in the example on the left. The vertical height of the waveform can be adjusted using the vertical sensitivity (volts/cm) control to vary the amplification factor. of the vertical amplifier. Vertical deflection is directly proportional to given input voltage levels and the sensitivity setting. The vertical sensitivity control is calibrated to insure accurate signal representation. For example, suppose the vertical sensitivity control was set to the 10 volts-per-centimeter position, as shown in the example. This would mean that each centimeter of vertical height or deflection would represent 10 volts at the vertical input terminals. The waveform is 4 centimeters high. Therefore, it has a peak-to-peak amplitude of 4 times 10 or 40 volts peak-to-peak. The peak value of the waveform would be equal to one-half of 40, or 20 volts peak. The value of any point along the waveform can be determined simply by comparing its relative position to the squares on the graticule and multiplying by the vertical sensitivity setting.

Measuring Period and Frequency

The oscilloscope may also be used to measure the period of an AC waveform. The period, or time for 1 cycle, is determined by observing the horizontal width of the waveform displayed on the screen. The oscilloscope's sawtooth oscillator can be adjusted so that the electron beam will move from left to right across the screen at a specific speed. The time required for the beam to move horizontally across the screen is referred to as the *sweep time.* The sweep time can be adjusted by the time base control. The time base sets the amount of time, in seconds, milliseconds, or microseconds, required for the trace to move horizontally a distance of 1 centimeter.

Assume that the oscilloscope's sweep time control is set to 5 ms/cm as shown in the figure on the right. This would mean that each centimeter of horizontal deflection would represent a time interval of 5 milliseconds. The waveform being displayed is 4 centimeters wide. In other words, one complete cycle occupies 4 centimeters of the trace. Therefore, the period of the waveform is equal to 4 times 5, or 20, milliseconds.





The sweep time control can be set so the oscilloscope can be used to measure waveforms that have very long and very short periods. You simply adjust the control to produce at least one complete cycle of the AC signal on the CRT. Then, count the number of centimeters from the beginning to the end of the cycle and multiply that number by the sweep time. Often, it is more convenient to use wave tops to determine period – any 2 consecutive wave tops is one period. Furthermore, many technicians will measure the period of a larger number of complete cycles then divide the measured period by the number of cycles to get the period of a single cycle.

In many cases, it is desirable to display only one cycle of a waveform, as shown. However, there are times when you need to observe more than one cycle, or a portion of a cycle. To show more than one cycle, you adjust the time base control for a shorter time period than you would use to display one cycle of the AC signal. For example, suppose you are measuring an AC signal with a 1 millisecond period and you set the sweep time control for 1 ms/cm. Then, you will see one complete cycle-per-centimeter across the CRT screen.

Naturally, to observe a portion of an AC cycle, you would set the sweep time control for a longer time period than you would use to display on cycle of the AC signal. In effect, you are magnifying a portion of the signal. For example, suppose the sweep time is 2 ms/cm and the CRT shows one complete cycle. Change the sweep time to 1 ms/cm and you will only see one-half of a cycle.

The frequency, *f*, of an AC waveform is easily calculated from the period, T, of the waveform. The frequency, in hertz, is expressed mathematically as:

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

For example, the waveform in the example trace above has a period of 20 milliseconds (0.02 seconds). The frequency (*f*) is, thus, 1/0.02 s = 50 Hz.

Investigation 10: Measuring AC Voltage and Frequency with the Oscilloscope

Materials Required

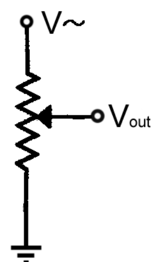
Trainer
Oscilloscope with 10x probe
Digital multimeter (DMM)
100 kΩ potentiometer (use the potentiometer on the Trainer)

Procedure

Record all of your data.

1. Set up the DMM to AC voltage.

2. Assemble a variable voltage divider circuit similar to the one in Investigations 8 and 9 using the 100 kΩ potentiometer:



3. Use the sine wave function generator on the Trainer for the AC supply voltage. If the function generator on the Trainer has its own potentiometer to adjust output voltage, set it to the maximum level for this investigation. If the function generator has an "offset" adjustment[8], set the offset to either zero volts or adjust the control about midway through its range.

[8] Many function generators have a DC offset control which allows for the adjustment of the DC component of the AC signal produced.

4. Set the function generator to produce a sine wave range of about 1000 Hz. Do not trust the frequency markings on the function generator – they are only approximate – since you will use the oscilloscope to measure the output frequency.

5. Turn on the oscilloscope. Attach the shielded probe lead to the BNC connector on the input to channel 1 and ensure that only channel 1 is active[9]. For now, set the time base to 1 ms/cm and the vertical sensitivity to 1 V/cm (these will be adjusted later). Set the trigger mode to AUTO and set the trigger level to zero. A switch or bank of buttons is usually provided to select the input "coupling" mode. Set the input coupling mode to AC[10]. There may be other settings which will need to be checked, as well – the instructor will specify other settings if necessary. Attach the ground clip to the breadboard ground and the test probe to V_{out}.

6. Turn on the Trainer. Adjust the potentiometer to point about midway through its travel. A sine wave should be visible on the oscilloscope. If not, check 1) the probe connections, 2) the trace "intensity" control to ensure the display is sufficiently bright, 3) the horizontal and vertical position controls (the trace may simply be moved off of the display), 4) the trigger level (the signal may be too weak to trigger the oscilloscope at the selected sensitivity) and 5) the input coupling mode to ensure that it is not set to ground (gnd) or DC.

7. Measure the AC peak-to-peak voltage with respect to ground. Adjust the potentiometer to show that you can control the AC voltage in a similar manner as DC voltage. Adjust the vertical sensitivity and time base controls on the oscilloscope to become comfortable with the effects observed by changing these settings.

8. Adjust the 100 kΩ potentiometer to give an output voltage (V_{out}) of about 500 $mV_{p\text{-}p}$. Adjust the vertical sensitivity to make the waveform display vertically across about $^2/_3$ of the graticule. The vertical scaling is not critical as long as the voltage can be read with suitable precision. Adjust the time base to display 5-10 complete AC cycles.

9. Change the input frequency and observe the effect on the oscilloscope.

10. Set the frequency control back to the 1000 Hz mark (remembering that the marks on the Trainer are only approximate). Measure the period of the wave using 2 consecutive wave tops. Measure the time required for, say, 5 wave tops and convert that value to period. Convert the measured periods to frequency.

11. Calculate the V_{rms} from the measured peak voltage. Attach one of the DMM leads to ground and the other to V_{out} and record the measured rms voltage from the DMM. Is the rms voltage measured by the DMM close to the rms voltage calculated from V_p? If not, the DMM frequency response is probably limited to a maximum of about 100 Hz. If so, then you can use the DMM for rms voltages up to at least 1000 Hz. How would you check for an even

[9] Most modern oscilloscopes have at least 2 channels which allow the technician to observe two different inputs simultaneously. This allows for immediate comparisons of voltage, phase, etc.

[10] Setting the coupling mode to AC inserts an internal capacitor in series with the probe to block any DC component the waveform might possess.

higher frequency response? Perform the experiment if time permits since this information will prove useful later.

12. Do not disassemble this circuit yet.

13. Set the function generator to produce triangle and square wave output. Record the difference in wave shape for these waveforms. Set the function generator back to sine wave for the next investigation.

Discussion

The utility of the oscilloscope for analyzing a simple AC circuit should be obvious from the simple measurements made in this investigation. If you explored the oscilloscope controls in some detail, you might have found that you could change the time base to display less than one complete cycle all the way to where you could not differentiate the hundreds or even thousands of wave tops. You may have also elucidated the maximum frequency response of the your DMM, an important parameter for this tool.

ELECTRICAL MEASUREMENTS VI

Frequency Counter

Oscilloscopes are extremely useful for observing repetitive and transient effects. They are also useful for determining the frequency of a waveform. However, you may have discovered in Investigation 10 that determining frequency of a wave on the oscilloscope is not fast and, depending upon the time base setting, may prove difficult to get an accurate result.

An important instrument in designing and troubleshooting AC circuits is the frequency counter. The frequency counter simply samples the oscillating signal and displays the frequency in hertz (or kHz, MHz, etc.). The frequency counter, while not a required instrument in the technician's toolkit, provides a level of convenience and speed welcome by anyone working with AC signals.

The frequency counter can be purchased as a stand-alone instrument; however, many DMMs now come with a frequency counter built in.

Investigation 11: Measuring AC Frequency with the Frequency Counter

Materials Required

Trainer
Oscilloscope with 10x probe
Frequency counter
100 kΩ potentiometer (use the potentiometer on the Trainer)

Procedure

Record all of your data.

1. Use the same circuit from Investigation 10. Attach the frequency counter probes to the circuit, one at ground and the other at V_{out}. Set the output voltage to at least 1 V_p but don't exceed 10 V_p.

2. Measure the frequency of the sine signal with the frequency counter as well as with the oscilloscope for 3 different frequencies (*e.g.*, 1000 Hz, 10 kHz, 100 kHz).

Which method for measuring frequency is most convenient. Which do you think is likely to be more accurate? Why?

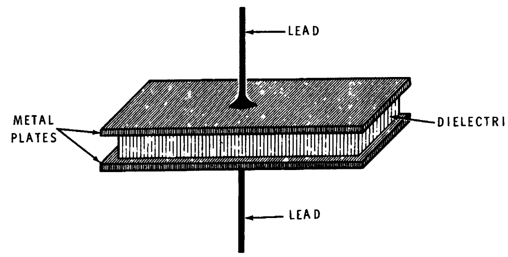
Discussion

The frequency counter indeed is faster and more convenient for quickly measuring the frequency of AC signals. Because of this, most technician's who frequently work with AC signals invest in a frequency counter.

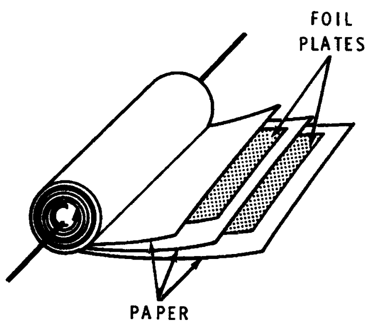
CAPACITANCE

The property of a component (or circuit) to store electrical energy in an electrostatic field is its *capacitance.* The component specifically designed for this purpose is the *capacitor.* The quantity of charge a particular capacitor can store is a measure of the device's capacitance.

Practically speaking, a capacitor is physically a very simple device consisting of two metallic plates separated by a *dielectric* (see the figure on the right). The dielectric is an insulating material can be nothing more than air or it can be constructed from materials such as paper, mica, glass, polymers, ceramics, or other insulating materials.



Simplified Diagram of a Capacitor



A real capacitor generally looks nothing like the preceding diagram. The very common electrolytic capacitor, for example, uses two thin sheets of metal foil separated by paper saturated with a light oil. Additional oiled papers sheets are placed on the top and bottom of the sandwich of foil and paper then rolled into a compact cylinder. One lead is art ached to each sheet and the entire unit sealed in a protective package. An example of an *axial lead* capacitor is shown on the left.

The storage capacity of a capacitor is related to the area of the two metal plates, the distance by the plates are separated, and the dielectric constant of the insulator. The formula which relates these values is

$$C = \frac{\varepsilon A \kappa}{d}$$

where A is the area of the metal plate (in m^2), d is the distance the plates are separated (in m), ε is the permittivity constant (8.85×10^{-12} F/m), and κ is the dielectric constant of the insulator. A table of dielectric constants for selected substances is shown on the right. Notice that high capacitance devices can be made with small plate areas if insulators with high dielectric constants are utilized. This means that high value capacitors can be produced in small packages if necessary.

Some typical dielectric constants for common insulators

Material	Dielectric Constant (κ)
Air/vacuum	1
Waxed paper	3-4
Mica	5-7
Glass	4-10
Rubber	2-3
Ceramics	10-5000

The unit of capacitance is the farad (in honor of Michael Faraday). One farad is the capacitance which will store one coulomb of charge at one volt applied electric field. One farad is, however, a very large capacitance. Typical capacitances used in electronics are on the order of picofarads to 1000s of microfarads.

A capacitor is said to be a reactive component; that is, unlike a resistor changes in the voltage applied to a capacitor are not accompanied by an immediate response. This delayed reactance will provide the basis for constructing, among other circuits, passive filters and oscillators.

RC Time Constants

When a capacitor is connected across a DC voltage source, it charges to the applied voltage. If the charged capacitor is then connected across a load, it will discharge through the load. The length of time required for a capacitor to charge or discharge can be computed if certain circuit values are known.

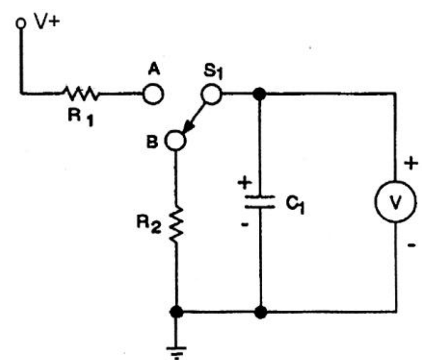
In the first approximation, only two factors determine the charge or discharge time of a capacitor: 1) the value of the capacitor and 2) the resistance through which the capacitor must charge or discharge. The next investigation explores this relationship.

Investigation 12: Measuring the RC Time Constant

Materials Required

Trainer
DMM
47 μF electrolytic capacitor (any electrolytic capacitor near this value is okay)
470 μF electrolytic capacitor (any electrolytic capacitor near this value is okay)
two 220 kΩ resistors (any two equivalent resistors from 100 kΩ to 470 kΩ are okay)
two 10 kΩ resistors
stopwatch

Procedure



1. Set up the DMM to read DC voltage.

2. Assemble the circuit shown. V+ = 10 V; $R_1 = R_2 = 220$ kΩ; $C_1 = 47$ μF. Electrolytic capacitors are polarized. Observe the correct polarity when connecting the capacitor into the circuit. S1 is a SPDT[11] switch. The center terminal of the switch is "common" and should be connected to the capacitor side of the circuit.

[11] SPDT = single-pole double-throw

3. Set the switch in the position which will discharge the capacitor through R_2. Wait for the DMM to read 0 V before proceeding.

For this part of the investigation, you will need to read the voltage drop across the capacitor every 5 seconds for at least 1 minute after you set the switch to the position which will cause the capacitor to charge. The voltage drop across the capacitor is a function of it level of charge, thus at time zero you can assume that the voltage drop will be 0 V.

4. Set the switch to the alternate position and start the stopwatch simultaneously. Record the voltage drops for at least a 1 minute interval. Notice that as time progresses the voltage across the capacitor stabilizes at or near supply voltage. Stop recording data when no further change in the voltage drop is observed.

5. Reset the stopwatch. Set the switch to the discharge position and record the voltage drop across the capacitor every 5 seconds until the voltage drop is 0 V. You can assume the voltage drop at time zero is the supply voltage.

6. Repeat the investigation with the 470 μF capacitor and 10 kΩ resistors. It may be necessary to record the voltage drops at more frequent intervals.

7. Make a graph of the voltage drops versus time in the charging and discharging direction for both resistor-capacitor combinations. You can combine the charge and discharge plots on the same graph.

Discussion

You may have observed that the charge and discharge plots appear to be exponential curves (in fact, they are). The RC time constant is defined as the time required for the capacitor to reach 63.2% of its full charge (or, conversely, to discharge by 63.2% of its full charge) in the RC circuit.

Examine the capacitor charging plot. Determine the time necessary to achieve 63.2% of full charge. If you used 10.0 V for source voltage, this is simply the time necessary to reach a voltage drop of 6.32 V. Do the same analysis for the discharge curve and repeat the analysis for the second RC circuit.

Can you deduce, by examination, the relationship of the values of R, C, and the time constant, *t?* Converting the resistance and capacitance values to their base unit value *(i.e.*, kΩ to Ω, μF to F) may help. Also, repeating the investigation with additional R-C combinations may further help.

If you collected the data carefully (and the component values are near to the stamped values), you probably discovered that the time constant, *t,* is related to the values of the resistor and capacitor by the simple mathematical relationship

$$t = R \times C$$

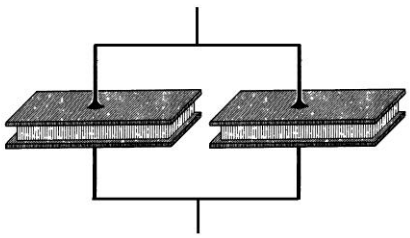
You may have also noticed that a completely discharged capacitor behaves as a wire when a DC voltage is initially applied to the device and as an open circuit when the capacitor is fully charged. This concept will become ever more important as you design circuits containing capacitors.

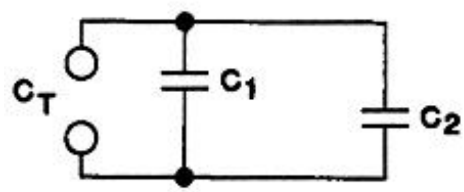
CAPACITOR NETWORKS

Capacitors are sometimes connected in parallel but rarely in series combinations. In order to analyze and understand electronic circuits, we must be able to compute the total capacitance of capacitance networks.

Capacitors In Parallel

In parallel capacitor network, components are connected across each other so that there are two or more paths for current flow. The net effect is that the overall plate area is increased.





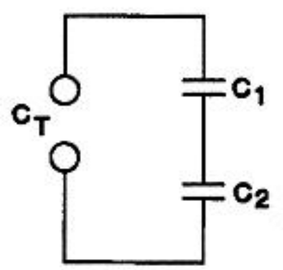
Since the effective plate area is increased, the total capacitance, C_T is easily calculated as

$$C_T=C_1+C_2+...$$

Capacitors are often placed in parallel to adjust or "trim" the value to a very specific capacitance. Variable trimmer capacitors are available which are often placed in parallel with fixed value capacitors for fine-tuning transmitters or receivers in radio frequency applications.

Capacitors In Series

Less frequently, capacitors are connected in series. The total capacitance of a series capacitor network

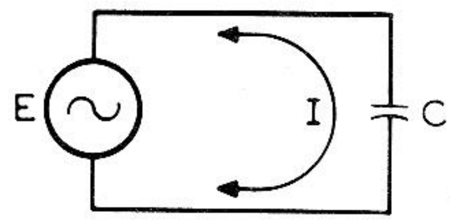


is analogous to a parallel resistor network. The total capacitance is calculated as

$$C_T = \left(\frac{1}{C_1} + \frac{1}{C_2} + \cdots + \frac{1}{C_n} \right)^{-1}$$

CAPACITORS IN AC CIRCUITS

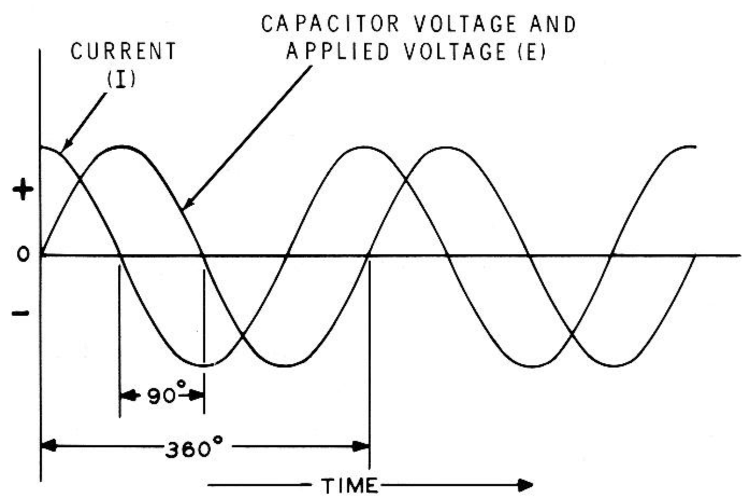
When an AC voltage is applied to a capacitor, alternating current flows in the circuit. As the AC voltage varies, the current in the circuit follows in the same way. While electrons do not pass from one plate to the other through the dielectric, electrons do flow in the circuit external to the capacitor as if they did. As the applied AC voltage rises and falls, the capacitor charges and discharges. This effect makes the capacitor appear to an AC source as a resistor with "resistance" dependent upon the frequency of the AC voltage and the value of the capacitor.



Current-Voltage Relationships in Capacitive AC Circuits

The relationship between the current and the applied voltage in a capacitive circuit is different from purely resistive AC circuits. In a circuit where AC voltage is applied to a resistance, current through the resistor follows the voltage applied to it. We say that the current and voltage in such a circuit are in phase. By virtue of the capacitor charging (increasing voltage drop) during the positive portion of the AC cycle and the capacitor charging with the opposite polarity during the negative portion of the cycle, the current and voltage in a capacitive AC circuit are not in phase with one another.

In a capacitive AC circuit, the capacitor constantly charges and discharges with a change in the applied voltage. Once the capacitor is initially charged, the voltage across it acts as a voltage source. Its effect is to oppose changes in the external supply voltage. Since the capacitor must charge or discharge to follow the changes in the applied voltage, the resulting current flow is out of step with the changes in the applied voltage. We say that there is a phase shift between the voltage and current in the circuit.



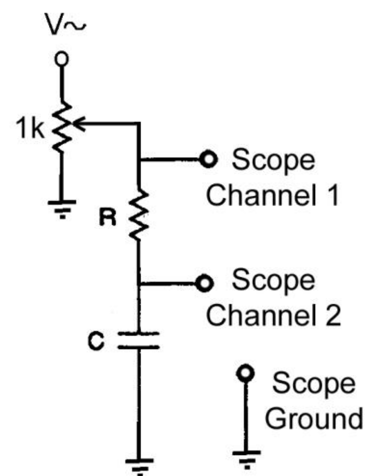
The exact relationship between the current and voltage in a capacitive circuit when a sine-wave AC signal is applied is shown in the figure on the right. Note that when the current is maximum, the voltage across the capacitor is zero. As you can see, there is a phase shift between the current and voltage in the circuit. This phase shift is expressed in terms of degrees. Recall that one complete cycle of a sine wave is 360 degrees. The phase shift in the purely capacitive circuit is 90 degrees. We say that the current and voltage in a purely capacitive circuit are 90 degrees out of phase with one another. Another important fact to note is that the change in current *leads* the change in voltage. Looking at the figure you can see that the capacitor voltage change *follows* the current change in time.

Investigation 13: Observing the Phase Shift in a Capacitive Circuit

Materials Required

Trainer
Two channel oscilloscope with two 10x probes
0.1 μF capacitor (Mylar and ceramic capacitors may be stamped 104 indicating 10×10^4 pF)
1 MΩ resistor
Frequency counter (optional)

Procedure



1. Assemble the circuit shown. As will be seen in later investigations, it is important to connect the capacitor to ground in this investigation. Use the l kΩ potentiometer on the Trainer to adjust the output voltage of the sine wave generator.

2. Connect channel 1 of the oscilloscope (some oscilloscopes refer to this as channel A) to the junction of the potentiometer and resistor. Hook up the ground lead to the circuit ground. Use AC coupling and AUTO trigger.

3. Set the waveform generator to "sine wave." While monitoring the frequency at the output of the potentiometer, adjust the frequency to about 1000 Hz. Adjust the oscilloscope time base to display a convenient number of cycles and the vertical sensitivity to conveniently display the waveform on the screen. Set the output voltage of the sine wave generator to about 1 V_p.

4. Connect channel 2 of the oscilloscope (some oscilloscopes refer to this as channel B) to the junction of the resistor and capacitor. No ground lead is necessary since the oscilloscope is already grounded by channel 1. Set the vertical sensitivity to the same as channel 1.

5. Observe and record the superimposed waveforms. Is the phase shift readily apparent? Change the frequency a few hundred hertz up and down. Does the phase shift change?

Discussion

This circuit is known as a low-pass filter. The phase of the voltage measured at the resistor (channel 1) is the same as the current phase since resistors do not introduce a phase shift in AC circuits. The voltage drop across the capacitor (channel 2) is approximately 90° phase shifted from the current through the capacitor. The phase shift is not exactly 90° since the circuit is not purely capacitive.

This investigation illustrates that two passive components can function together to perform a useful task, that is, filtering and/or phase shifting.

CAPACITIVE REACTANCE

By now, you realize that a capacitor is essentially an infinitely high resistance to DC current. You might also have assumed that the converse is also true; namely, that a capacitor is a pure conductor to AC current. You would, however, be only partially correct in the latter assumption. In an AC circuit, the voltage stored on the capacitor is always in constant opposition to the applied voltage making it appear to have resistance to AC current flow. This resistance to AC current flow is call *capacitive reactance,* X_C, and like resistance has the units of ohms.

Capacitive reactance is a function of the capacitance of the device and frequency of the applied AC voltage. High values of capacitance yield low reactance at a given frequency and present a low reactance to high frequencies. The capacitive reactance is calculated by

$$X_c = \frac{1}{2\pi f C}$$

where X_C is the capacitive reactance in ohms, f is the frequency in hertz, and C is the capacitance in F. Notice that the capacitive reactance is ∞ at DC (f = 0 Hz).

Reactance is just a generalized form of resistance. The AC reactance of a resistor is the same as the DC resistance so we do not make a distinction between the two. The AC reactance of a capacitor increases as frequency decreases.

For purely capacitive circuits (rare, to be sure), Ohm's law is obeyed with the modification that X_C is substituted for R:

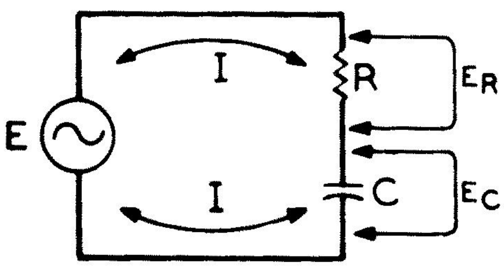
$$E = I \times X_c$$

Unfortunately, in AC circuits utilizing RC networks it is not possible to simply sum the DC resistance of the resistor with the AC reactance of the capacitor to get total AC resistance since there are differing phase relationships of the two devices. Since these phase differences exist, it is necessary to use a vector approach to add the quantities of resistance and reactance.

RC CIRCUITS

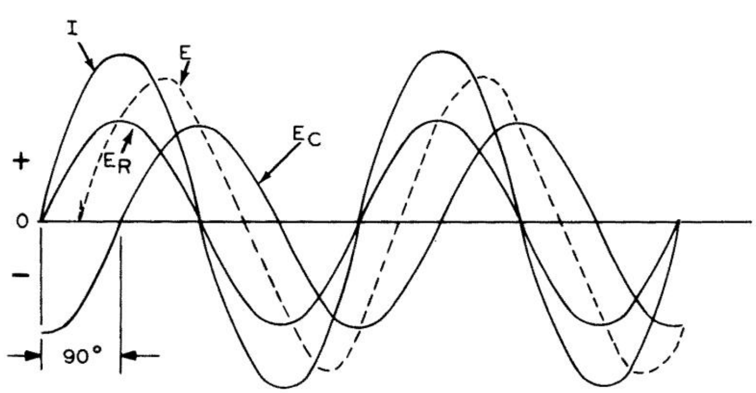
While purely capacitive circuits are occasionally used in electronic circuit, more often capacitors are combined with other components to form electronic networks. Among the most commonly used circuit is a series resistor and capacitor. Despite its simplicity, this simple series RC circuit has many applications. Another commonly used capacitor circuit is the parallel resistor-capacitor combination. While not as common as the series RC circuit, the parallel RC circuit is frequently found in electronic equipment. This section investigates the characteristics of both series and parallel RC circuits.

Series RC Circuits



The simplest form of RC circuit is a single capacitor connected in series with a resistor (as in Investigation 14). The figure on the left shows such a circuit connected to a source of AC voltage. The source voltage generates a sine wave of rms voltage E. The applied voltage will cause current to flow in the circuit. The capacitor will charge and discharge as the AC voltage varies. No electrons pass through the capacitor, but the charging and discharging action causes a movement of electrons in the circuit. The charge on the capacitor causes a voltage to be developed across it. This voltage is designated E_C. The voltage across the capacitor is a sine wave that lags the current flowing in the circuit by 90° (or, said another way, the current leads the voltage across the capacitor by 90°).

The current in the circuit passes through resistor R and develops a voltage drop across it. The voltage across the resistor, E_R, is in phase with the current in the circuit. The voltage across the resistor is a function of the resistance and the current ($E_R = IR$). The voltage across the capacitor is determined in the same way. The capacitor voltage drop is a function of the current flowing in the circuit and the capacitive reactance or $E_C = IX_C$.



The preceding figure shows the phase relationships of the various voltages and currents in this series RC circuit. In this illustration, the current sine wave, I, is used as the reference since one of the fundamental characteristics of a series circuit is that the same current is common to all components. Note the voltage drop across the resistor E_R. This voltage drop is in phase with the current. Now refer to the capacitive voltage drop E_C. The current through a capacitor leads the voltage across it by 90°.

An important characteristic of capacitive AC circuits is that Kirchhoff's laws still apply. Kirchhoff's voltage law tells us that the sum of the voltage drops across the components in a series circuit equals the applied voltage. This law is valid for the RC circuit when phase relationships are considered. If we add E_R and E_C by summing the amplitudes of the two sine waves at multiple points and plotting the resulting curve, we will obtain the sine wave represented by the dashed line. This waveform represents the applied voltage, E. It is important to note is that the applied voltage is not in phase with the current or with either the capacitor or resistor voltages. The current in the circuit leads the applied voltage as it will in any capacitive circuit but since the circuit is not purely capacitive, the current leads the applied voltage by some angle less than 90°. In the example, the current leads the applied voltage by approximately 45°. This difference in phase is called the phase shift or, more commonly, the phase angle.

Summary

If the circuit had no DC resistance (*i.e.*, if it was purely capacitive) then the current through the circuit would lead the applied voltage by 90°. On the other hand, if the circuit was purely resistive, the current and voltage would be exactly in phase. With both resistance and capacitance in the circuit, the phase difference between the current and the applied voltage is some value between 0° and 90°. The exact amount of the phase shift is determined by the amount of resistance and capacitive reactance in the circuit. Before we get into phase angle calculations, however, let’s take a closer look at the voltage relationships in a series RC circuit.

KIRCHHOFF'S VOLTAGE LAW IN SERIES RC NETWORKS

When using Kirchhoff’s law to analyze a series RC circuit, you must realize that it is not possible to directly add the numerical values of the voltage drops across the resistive and capacitive portions of the circuit to obtain the applied voltage. The voltage drops across the resistance and the capacitive reactance are not in phase with one another. In order to obtain the correct applied voltage, it is necessary to correct for the phase difference between the two voltages. You can do this with vector addition.

Vector Diagrams

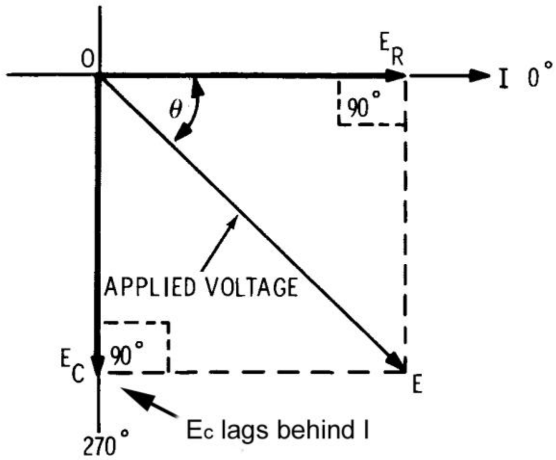
In the physical world, there are quantities that are expressed in specific units. Ohms, centimeters, and grams are all examples of these quantities. These quantities are measured in scalar values and in order to find the sum of a number of like quantities you would just add them together. For example, the total resistance in a series resistor network consisting of a 100Ω resistor and 220Ω resistor is

$$100\Omega + 220\Omega = 320\Omega$$

Problem solving using scalar quantities is easy because it involves simple addition.

Some physical quantities – vector quantities – have two properties; magnitude and direction. Velocity of an aircraft along a compass direction, force of gravity, and voltage at 90° are all examples of vector quantities.

It is possible to add vectors but, when you are doing this, you must take into consideration both the magnitude and the direction of the vector. In order to make this easier, it is often convenient to draw a vector diagram when you are analyzing an AC circuit.



The figure on the left shows the vector diagram for the voltages in a series RC circuit. A current vector, I, is shown on the x axis pointing to the right. (0°) This vector represents the value of the current flowing in the series RC circuit. It is used as a reference vector for the diagram because the current value is the same at all points in a series circuit. Coinciding with this vector is another vector labeled E_R. The length of this vector, from the origin to the point of the vector, represents the voltage dropped across the resistive portion of the circuit. This is the rms voltage as it would be measured with a voltmeter. The voltage vector overlaps the current vector because these two are in phase.

The voltage drop across the capacitor is labeled E_C. Again, this is the rms voltage measured across the capacitor. Notice that the direction of the capacitor voltage is shifted 90° from the direction of the resistor voltage.

The applied voltage, E, is the vector sum of the capacitor and resistor voltages, shown graphically in the figure above.

The angle formed by the applied voltage vector and the resistive voltage vector represents the phase shift between circuit current and circuit voltage. This angle is always between 0° and 90°. Later, we will calculate the phase angle but for now let's get back to vector addition. Symbolically, the addition of vector quantities is

$$\vec{E} = \vec{E}_R + \vec{E}_C$$

But this kind of vector addition contains more information than is normally necessary for the technician. The scalar magnitudes of the voltages are all that are normally required.

Since the applied voltage forms the hypotenuse of a right triangle formed from E_R and E_C, the scalar magnitude of the applied voltage is

$$\mathrm{E} = \sqrt{\mathrm{E}_\mathrm{R}^2 + \mathrm{E}_\mathrm{C}^2}$$

Example

What is the voltage (V_{rms}) drop across the capacitor in a series RC circuit where the applied voltage is 10 V_{rms} and resistor voltage drop is 6 V_{rms}.?

$$E = \sqrt{E_R^2 + E_C^2}$$
$$10\ V = \sqrt{6^2 + E_C^2}$$
$$E_C^2 = 100 - 36 = 64\ V^2$$
$$E_C^2 = 8\ V_{rms}$$

IMPEDANCE

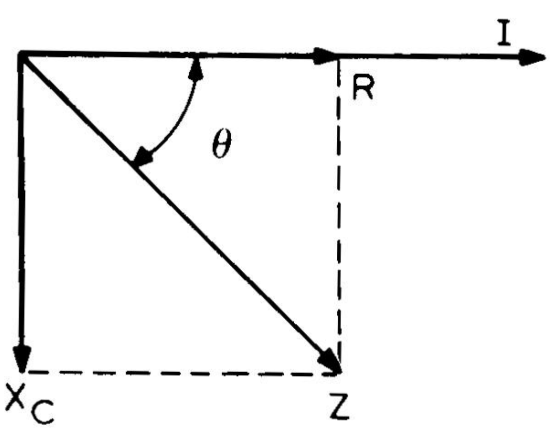
Impedance is the total opposition to current flow in any circuit although the term is often reserved for the apparent resistance in an AC circuit. In short, impedance is generalized resistance of a circuit including contributions from resistors, capacitors, inductors, etc. In a circuit consisting of a resistor and a capacitor, the total opposition is the sum of the capacitive reactance and the DC resistance. Both the reactance and the resistance impede current flow. For an RC circuit, the impedance is the vector sum of the capacitive reactance and resistance.

The impedance of an AC circuit is expressed in ohms and is designated by the letter Z. We can define the impedance in terms of Ohm's law just as we defined the total resistance of a DC circuit.

$$E = I \times Z$$

This expression can be rearranged using basic algebra to obtain the expressions for voltage and current in terms of the circuit impedance:

In the previous section, we saw that because of the phase shift caused by the capacitor in a series RC circuit, the voltage drops across the capacitor and resistor could not be added directly to obtain the applied voltage. Instead, a vector sum had to be taken in order to obtain the correct value. Since the current through a series circuit is the same in all elements, we can say that the voltage drops across the circuit components are directly proportional to their resistance or reactance. For that reason, we can draw a diagram exactly like the voltage vector diagram described earlier, to obtain the total impedance of the circuit, as shown on the left. Here the current vector is again used as the reference. The resistance vector is coincident with the current vector since the resistive voltage drop is in phase with the current. In this case, the length of the vector is proportional to the resistance. Another vector representing the magnitude of the capacitive reactance, X_C, is drawn 90° out of phase with the resistance vector to take into account the 90° phase shift produced by the



capacitor. The impedance is the magnitude of the hypotenuse of the right triangle formed:

$$Z = \sqrt{R^2 + X_C^2}$$

Example

What is the impedance of the series RC circuit at 1000 Hz consisting of a 1.0 kΩ resistor and a 0.1 μF capacitor?

First, the capacitive reactance must be calculated at 1000 Hz. Using the equation

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

the capacitive reactance is calculated as

$$X_C = \frac{1}{2\pi(1000\text{Hz})(0.1\times10^{-6}\text{F})} = 1592\Omega$$

The impedance is, then,

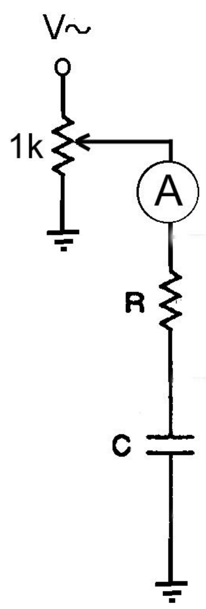
$$Z = \sqrt{1000^2 + 1592^2} = 1880\Omega$$

Investigation 14: Validating Ohm's Law in a RC Circuit

Materials Required

Trainer
DMM
0.47 μF capacitor (0.22-0.5 μF)
1 kΩ resistor

Procedure



1. Assemble the circuit shown. Use the 1 kΩ potentiometer on the Trainer to adjust the output voltage of the 60 Hz 15 VAC supply on the Trainer (do not use the "floating 30 VAC supply). Using the DMM to read AC voltage, adjust the output of the potentiometer to 10 VAC.
2. Set the DMM to measure AC current and connect it into the circuit between the potentiometer and resistor.
3. Measure and record the current.
3. Remove the ammeter from the circuit and connect the output of the potentiometer directly to the resistor. Measure and record the AC voltage drop across R and C. **Remember to change the DMM to read voltage to prevent damage to the DMM.**

Calculate the impedance of this circuit. From the impedance and applied voltage, calculate the current. How do the measured and calculated currents compare?

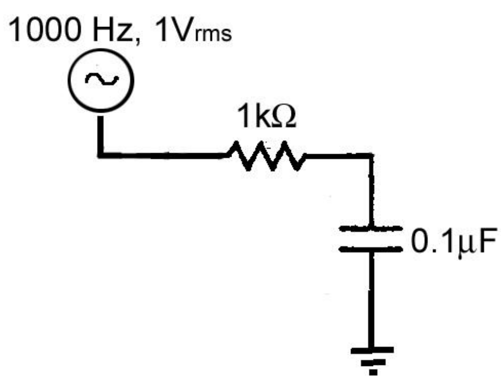
Calculate the sum of the voltage drops across R and C. Do they add up to the applied voltage? If not, why not? How should the voltages be summed?

PHASE SHIFT AND PHASE ANGLE

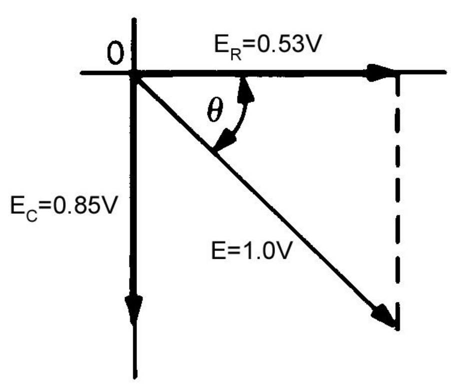
Knowing the phase shift in a RC circuit is often important when designing AC circuits, especially those involving oscillators, mixers, and radio frequency circuits. Computing the phase shift (also called the *phase angle)* is simple using the trigonometric functions.

The phase shift for a RC circuit can be calculated from either the voltage drops across the resistor and capacitor or from the resistance and capacitive reactance. The choice of which to use is determined only by convenience.

Consider the following circuit:



Careful measurement of the voltage drops across each component yields E_R = 0.53V and E_C = 0.85 V. Note that since this is a "reactive" AC circuit (*i.e.,* the circuit contains a reactive element, the capacitor) that the numerical sum of the voltage drops is *not* equal to source voltage. Draw the voltage vector diagram for this system. E_R is along the current axis and E_C is 90° phase shifted.



Observe that the magnitude of the vector sum of the voltage drops *is* the source voltage.

Using the trigonometric tangent relationship, the phase shift phase angle, θ) can be calculated:

$$\tan\theta = \frac{0.85}{0.53} = 1.6 \quad \text{so} \quad \theta = \arctan(1.6) = 58°$$

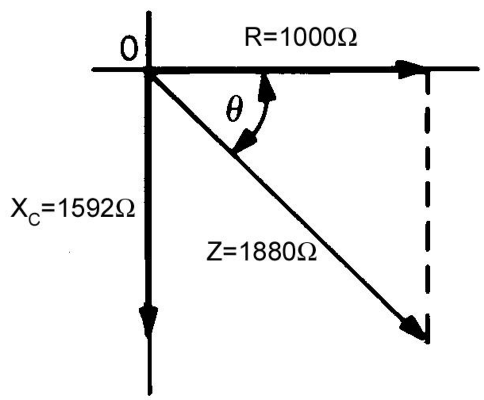
What this indicates is that this circuit shifts the phase of the 1000 Hz input signal by 58°.

An alternative approach – and more useful when designing circuits – to determine the phase angle is to use the resistance and capacitive reactance and trigonometrically calculate the phase angle in a similar manner as in the previous example.

Start by calculating the capacitive reactance of the capacitor at 1000 Hz:

$$X_C = \frac{1}{2\pi(1000\ \text{Hz})(0.1\times10^{-6}\,\text{F})} = 1592\Omega$$

Now, draw the vector diagram using R and X_C:



The phase angle, θ, is calculated using the tangent rule:

$$\tan\theta = \frac{1592}{1000} = 1.592 \quad \text{so} \quad \theta = \arctan(1.592) = 58°$$

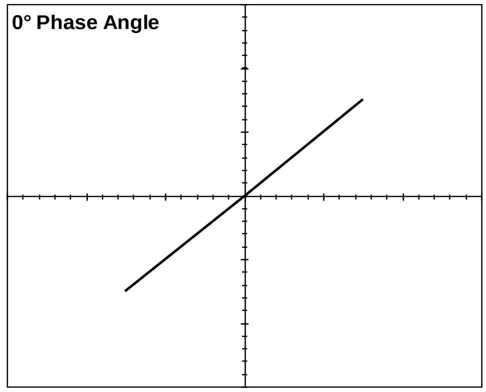
Of course, in both examples the sine or cosine rules could be used to arrive at the same result.

Investigation 15: Measuring Phase Angle in a Series RC circuit

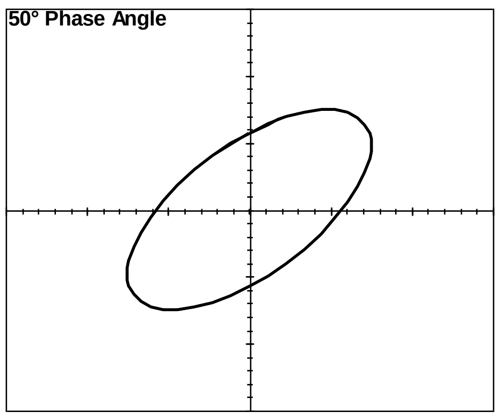
Introduction

The oscilloscope can be used to not only show the presence of a phase shift in a RC circuit (Investigation 14) but, properly setup, can also help to easily measure the phase angle graphically. The method initially appears complicated but is really rather simple in both setting up the instrument and analyzing the data.

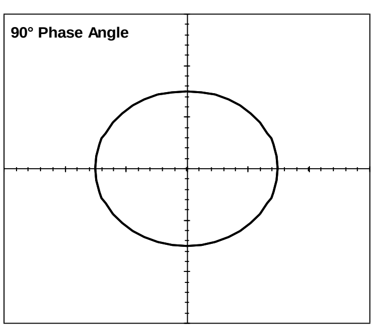
The technique takes advantage of a function found on most oscilloscopes; that is, the ability to substitute an external frequency supplied by the technician in place of the internal time base generator. In this way, the user supplied AC signal sweeps the display beam left and right while the test signal moves the beam up and down (in the normal way). If the two signals are perfectly in phase, a diagonal line with a positive slope is observed on the oscilloscope screen.

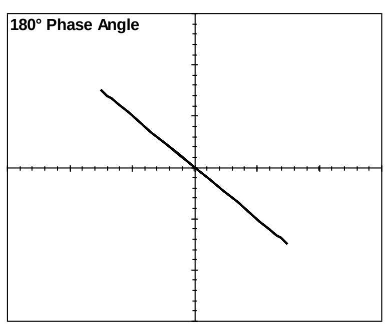


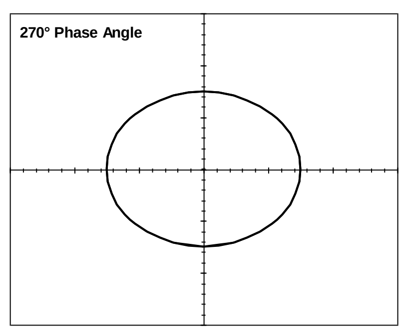
However, as the phase of the two signals starts to diverge a trace known as a Lissajous pattern forms. When the phase difference is, for example, 50° the following trace is observed:



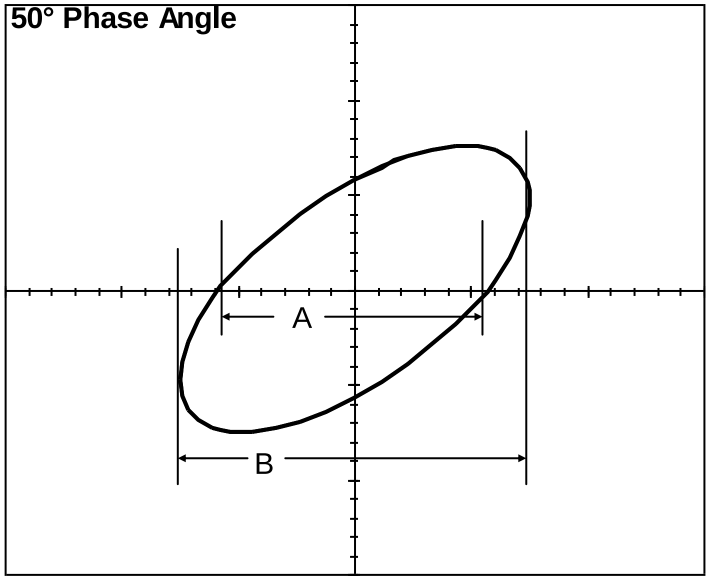
The following display images illustrate the view with increasing differences in phase angle:







But how is the Lissajous pattern used to measure phase angle? Without proof, the calculation is a very simple trigonometric relationship. Look at the following image of the 50° phase shifted signal with new annotations:



If the Lissajous pattern is properly centered on the display; the ratio of the x-axis contact-points (A) to the total width of the ellipse (B) is the sine of the phase angle. In this case, A = 11.5 units and B = 15 units. Thus,

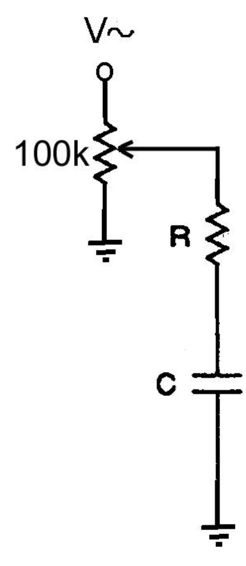
$$\sin\theta = \frac{11.5}{15} = 0.767$$

So, the phase angle, θ, is arcsin(0.767) or 50.1°. In this example, the RC circuit acts as a phase shifter which imposes a 50.1° phase shift between the input signal and output signal.

Materials Required

Trainer
Oscilloscope (with two 10x probes)
Frequency counter (optional but useful)
0.47 μF capacitor (0.22 – 0.5 μF)
100 kΩ resistor

Procedure



1. Assemble the circuit shown. This is similar to the circuit used in Investigation 14. Use the 100 kΩ potentiometer on the Trainer to adjust the output voltage of the sine wave generator. Do not use the 15 or 30 VAC terminals for this investigation.

2. Use an oscilloscope or frequency counter to set the frequency of the sine wave generator to 1 kHz. Use the oscilloscope to measure voltage and set the output of the potentiometer to 1 $V_{p\text{-}p}$ Do not use the AC voltmeter for this measurement, it may not have sufficient bandwidth for an accurate determination.

3. With channel 1 of the oscilloscope attached to the resistor and channel 2 attached to the capacitor, verify that the circuit is a phase shifter (see Investigation 14).

4. Leave channel 1 attached to the resistor and make channel 2 an external horizontal input. Various oscilloscopes use different ways to make this connection. The two most common ways are 1) an external horizontal input (EXT HORIZ) connector different from the channels 1 or 2 inputs with an accompanying switch to activate the input and 2) a panel switch to automatically convert one of the input channels to a horizontal input (X/Y setting is common).

5. Turn on the Trainer (if not already). Adjust the amplitude controls on the oscilloscope to get a Lissajous pattern of sufficient size to measure its characteristics. Make sure it is perfectly centered on the screen in the x and y direction.

6. Measure and record the parameters A and B as described in the Introduction. Calculate the phase shift produced by this circuit. Using R and X_C, calculate the theoretical phase shift which should be produced by this circuit at this frequency.

7. Slowly increase and decrease the frequency by about 1000 Hz and observe the change in phase shift. Which component is causing the phase shift to change (resistor or capacitor)? Does the phase shift change in the predicted direction with changes in frequency?

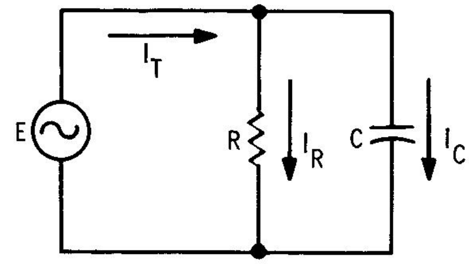
8. For added exploration, you may want to change the fixed resistor to the l kΩ potentiometer, using it as a rheostat; that is, using only terminal 1 and 2 (or 2 and 3). Experiment with the effect resistance on the phase shift. Try the experiment at different frequencies.

Discussion

The results that you obtained by trigonometric calculation (based on R and X_C) and direct measurement with the oscilloscope should be similar but may differ significantly. The accuracy of the calculation is determined by the accuracy of the stamped component values. While the resistors you use are probably ±5%, with many capacitors the stamped values may have tolerances exceeding ±10%. Nonetheless, it is clear that a series RC circuit acts as a phase shifter in a predictable manner.

The phase shifter will be vitally important in the design of feedback oscillators – devices used in a dazzling variety of electronic instruments and consumer electronics.

Parallel RC Circuits



Thus far, we have limited our discussion of RC circuits to the series RC circuit. It is also possible to connect RC circuits in parallel configurations. An elementary parallel RC circuit is shown on the right. As with the series RC circuit, Ohm's law applies to all of the components within the circuit and, when you determine circuit unknowns, it is necessary to use vector addition. However, when drawing vector diagrams for parallel circuits, there is one major difference.

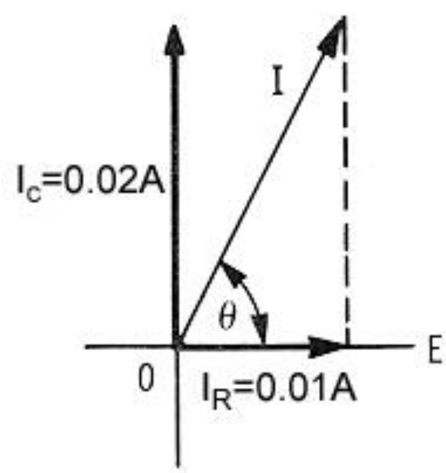
Recall, in a series circuit, current is used as a reference because it is the same at all points in the circuit. In a parallel RC circuit vector diagram, however, voltage must be used as the reference. Remember also the voltage drops across all branches of a parallel circuit are equal. Let's look at a basic circuit and some vector diagrams to see exactly what differences there are.

Consider the elementary RC parallel circuit above with and applied voltage of 1 V_{rms} at 60 Hz, R = 100 Ω, and X_C = 50 Ω. In this circuit, the voltage drop across both the resistor and the capacitor is 1 V. Since the voltage drop, as well as the capacitive reactance and resistance is known, it is possible to use Ohm's law to determine the current through the components:

$$I_{resistor} = \frac{1\ V}{100\ \Omega} = 0.01\ A = 10\ mA$$

$$I_{capacitor} = \frac{1\ V}{50\ \Omega} = 0.02\ A = 20\ mA$$

The vector diagram, using these two values, is shown on the right. Notice that the capacitor current leads the resistor current (opposite of the voltages in the series RC vector diagram). The total AC current flowing in the circuit is, thus,



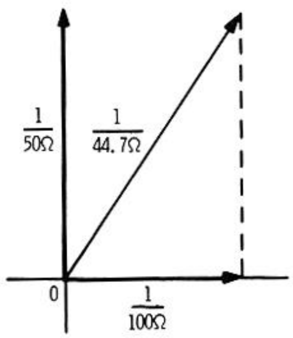
$$\begin{aligned} I &= \sqrt{I_R^2 + I_C^2} \\ &= \sqrt{0.01^2 + 0.02^2} \\ &= 0.022\ A \end{aligned}$$

The phase shift is calculated as before,

$$\theta = \arctan\left(\frac{I_C}{I_R}\right)$$

One last comment about parallel RC circuits. The impedance of the parallel RC network is calculated slightly differently. Recall in DC electronics that the resistance of a parallel circuit is always less than the smallest resistive branch in the circuit. The same general principal holds true for impedance in a parallel RC circuit. The impedance is always less than the smaller of the two current opposing components.

For this reason, when the impedance vector diagram is c6nstructed, you must use the reciprocal of the resistance and capacitive reactance, and the resulting vector will be the reciprocal of the actual impedance. This makes physical visualization of the effect of the vectors somewhat less intuitive but the math no more difficult.



The diagram illustrates this point. Notice the total impedance is less than either the capacitive reactance or the resistance. In the parallel circuit, impedance must be calculated using reciprocal values. Likewise, phase shift must be calculated from the reciprocal values.

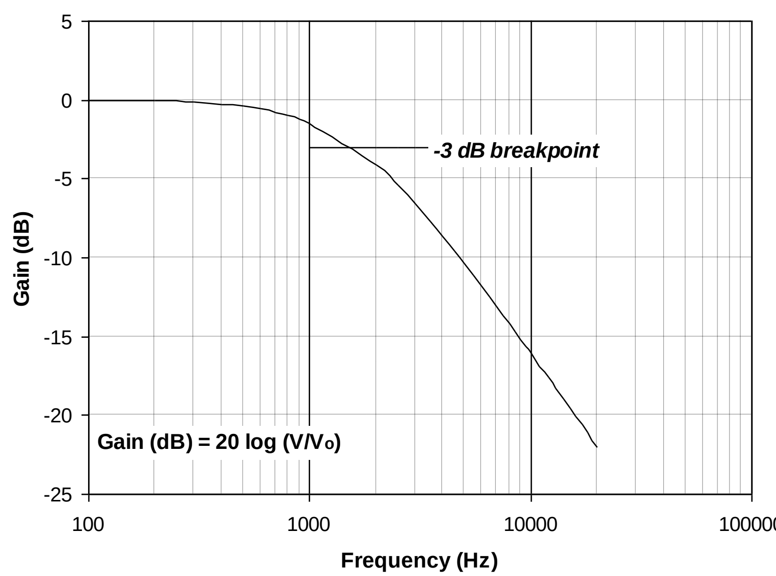
An alternative method to calculating the impedance of the network is to use an AC derivation of the parallel resistance formula:

$$Z = \frac{RX_C}{\sqrt{R^2 + X_C^2}}$$

APPLICATIONS OF RC CIRCUITS
LOW-PASS AND HIGH-PASS FILTERS

RC networks form the basis for a wide variety of practical circuits ranging from capacitive voltage dividers and coupler/decouplers to phase shifters and filters, just to name a few. In this section, we will restrict the discussion to one of the most common uses of the RC network – low-pass and high-pass RC filters.

A filter is a frequency discriminating circuit; that is, it strongly attenuates some frequencies while allowing others to pass with only slight attenuation. Three of the most common types of filters used are the low-pass, high-pass, and band-pass filters. The low-pass filter permits low frequency signals to pass from the input to the output with little attenuation while the high-pass filter passes high frequency and attenuates low frequency. The band pass filter allows a region of a broad band of frequencies to pass while attenuating both high and low frequencies.



The plot shown above is known as a Bode diagram (or, frequency response curve) and shows some of the features of a low-pass filter with a cut-off frequency (or, breakpoint frequency, f_b) of slightly over 1000 Hz. The gain of the filter is always less than zero since the filter will attenuate even the pass band a small amount. The voltage gain., A_v, is simply

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

Often, gain is reported in the units of *decibels* (dB). The *bel* (after Alexander Graham Bell) is the log of the ratio of *output power* to *input power* (rather than output voltage to input voltage) of any circuit. However, the bel is a rather large quantity so, more often, the decibel is used and reported. Gain, in dB, is calculated

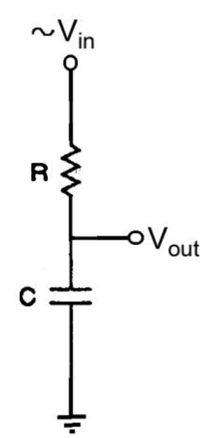
$$\text{Gain (dB)} = 10 \times \log \frac{P_{out}}{P_{in}} = 20 \times \log \frac{V_{out}}{V_{in}}$$

A negative gain (in dB) means that the circuit is a signal attenuator.

The breakpoint frequency of a filter is defined as the frequency where the gain is -3 dB; that is, the frequency where $V_{out} = 0.707V_{in}$. The technician will find it relatively easy to design RC filters with desired a cut-off frequency.

Low-Pass Filter

The simplest form of the low pass filter is shown on the right. This is the familiar circuit used in prior investigations. The output of the filter is taken between the capacitor and ground. The easiest way to understand the operation of the filter is to consider the circuit as an AC voltage divider with one reactive element (the capacitor).



Recall that the output voltage of a purely resistive voltage divider is

$$V_{out} = V_{in}\left(\frac{R_2}{R_1 + R_2}\right)$$

Substituting capacitive reactance and impedance (since this is an AC circuit) yields

$$V_{out} = V_{in}\left(\frac{X_C}{\sqrt{R^2 + X_C^2}}\right)$$

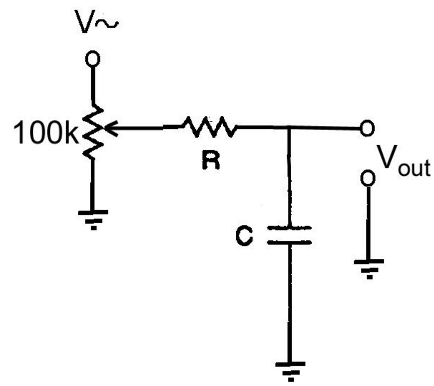
The value of R remains constant in an AC or DC circuit but X_C varies with frequency. At low frequency, the capacitive reactance is very high, thus most of the voltage is dropped across the capacitor and the gain of the circuit is near unity. As the frequency increases, the capacitive reactance decreases and the voltage drop across the capacitor decreases. At high frequency, where the capacitive reactance is very low, very little voltage drop is observed at the capacitor.

Investigation 16: Frequency Response of a RC Low-Pass Filter

Materials Required

Trainer
Oscilloscope (with two 10x probes)
Frequency counter (optional but useful)
0.1 μF capacitor (0.047 – 0.5 μF)
1 kΩ resistor

Procedure



1. Assemble the circuit shown. Notice that this is simply a series RC circuit similar to those used in Investigations 14 and 15. Use the l00 kΩ potentiometer on the Trainer to adjust the output voltage of the sine wave generator.

2. Use the oscilloscope to measure voltage and set the output of the potentiometer to 2 V_{p-p} (this will be V_{in} during the analysis). Use the oscilloscope or frequency counter to measure frequency and set the output of the sine wave generator to 100 Hz (or the lowest frequency obtainable).

3. Attach the oscilloscope channel 2 to the output of the filter. Remember, no additional ground is necessary if the channel 1 probe provides the instrument ground. Measure and record the input and output voltage of the filter.

4. Increase the frequency by about 100 Hz (the exact amount is not critical). Determine the frequency and verify that the input voltage has not varied. If necessary, readjust the input voltage. Measure and record the input and output voltage.

5. Repeat procedure 4 in steps of 100 Hz up to a final frequency of 1000 Hz, then increase the frequency in steps of 1000 Hz to a final frequency of 10,000 Hz.

6. Calculate the gain (in dB) of the filter and construct a Bode plot. Determine the -3 dB cutoff frequency. Since you are measuring voltage and not power, use the correct for of the gain equation.

Discussion

The cutoff (breakpoint) frequency for an RC filter is calculated as

$$f_b = \frac{1}{2\pi RC}$$

Compare the breakpoint frequency obtained experimentally with that which i& calculated from component values.

What is the relationship of R and X_C at the breakpoint frequency?

High-Pass Filter

The high-pass filter is similar to the low-pass filter in that it is a series RC circuit. However; by changing the order of the components in the circuit, it becomes a high-pass circuit. The high-pass RC filter can be analyzed in the same way as the low-pass filter – a voltage divider – except the output voltage is taken as the voltage drop across the resistor in this case.

The voltage dr6p across R is calculated

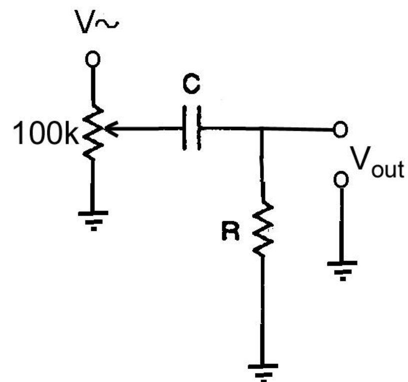
$$V_{out} = V_{in}\left(\frac{R}{\sqrt{R^2 + X_C^2}}\right)$$

At low frequency, the capacitive reactance is large thus imposing a large impedance to current flow. As the frequency increases, the capacitive reactance decreases and the gain of the circuit (the term in the parenthesis) approaches unity. The breakpoint frequency is calculated in exactly the same way as in the low-pass filter.

Investigation 17: Frequency Response of a RC High-Pass Filter

Materials Required

Trainer
Oscilloscope (with two 10x probes)
Frequency counter (optional but useful)
0.1 μF capacitor (0.047 – 0.5 μF)
1 kΩ resistor



Procedure

1. Assemble the circuit shown. If the circuit from Investigation 16 is still assembled, you can simply swap the resistor and capacitor. Use the 100 kΩ potentiometer on the Trainer to adjust the output voltage of the sine wave generator.

2. Repeat the procedure of Investigation 16.

3. Calculate the gain (dB) of the filter and construct a Bode plot. Determine the -3 dB cutoff frequency. Compare the experimental Bode plot to the theoretically derived plot.

Discussion

Compare the breakpoint frequency obtained experimentally with that which is calculated from component values.

If the component values used in Investigation 16 and 17 are the same, compare the breakpoint frequency of the two filters. Are they the same? Should they be the same?

Build other low- and high-pass filters using different component values. Convince yourself that the characteristics of the filter behave predictably.

For an even better understanding of the RC filter, measure and record the phase angle of the filter (see Investigation 15) over the entire frequency range used in the characterization of the low- or high-pass filter. Make a plot of phase angle, θ, versus log frequency. What is the significance of the phase angle at the breakpoint frequency?

DESIGN CHALLENGE - THE RC BAND-PASS FILTER

\

Design, build, and analyze a band-pass filter using a low-pass filter coupled to a high-pass filter. Design the filter such that the pass-band centers at approximately 1100 Hz with a low frequency breakpoint of~300 Hz and a high frequency breakpoint of~1400 Hz. Consider these parameters only starting points to get nominal component values. Remember that there are a limited number of common component values.

SEMICONDUCTORS

Semiconducting materials serve as the basic building materials used to construct some very important and indispensable electronic components. These semiconductor components are in turn used to construct virtually every useful electronic circuit. The three most commonly used semiconductor devices are diodes, transistors, and integrated circuits. However, other special components are also available.

The primary function of semiconductor devices in electronic equipment is to control currents or voltages in such a way as to produce a desired end result. For example, a diode can be used as a rectifier to produce pulsating DC from AC. A transistor can be used as a variable resistance to control the current in a load or to pr ovide voltage or current amplification. Integrated circuits can be constructed which contain tens to thousands of individual components in a single small package which are used in amplifiers, computers, and radio frequency electronics, just to name a few. All of these components are made of specially processed semiconductor materials such as silicon and germanium.

Semiconductor devices are extremely small, lightweight components which consume only a modest amount of power and are highly efficient and reliable. The vacuum tubes that were once widely used in practically all types of electronic equipment have been almost entirely replaced by semiconductor devices.

To be sure, one specific type of vacuum tube is still in ubiquitous use: the cathode ray tube used in televisions, computer display monitors, and oscilloscopes. However, even in this arena, the liquid crystal display (such as that used in laptop computer monitors and flat-screen display monitors) is making a significant impact in the phasing out of the vacuum tube.

Advantages

Components which are made of semiconductor materials are often referred to as *semiconductor* components because they are made from solid materials. Semiconductor devices are inexpensive, rugged and reliable, and are able to operate under extremely hazardous environmental conditions.

The nature of semiconductor devices also eliminates the need for filaments or heaters as found in all vacuum tubes. This means that additional power is not required to operate the filaments and component operation is cooler and more efficient. By eliminating the filaments, a prime source of trouble is also avoided because the filaments generally have a limited life expectancy. The absence of filaments also means that a warm-up period is not required before the device can operate properly.

Semiconductor components are also able to operate with very low voltages (between 1 and 25 volts) while vacuum tubes usually require an operating voltage of 100 volts or more. In fact, the television tube requires at least 10,000 V to operate the electron gun. Thus, semiconductor components generally use less power than vacuum tubes and are, therefore, more suitable for use in portable equipment which obtains its power from batteries. The lower voltages are also much safer to work with.

The most sophisticated semiconductor devices are integrated circuits. These are complete circuits where all of the components are constructed with semiconductor materials in a single microminiature package. These devices not only replace individual electronic circuits but also complete pieces of equipment or entire systems. Entire computers and radio receivers can be constructed as a single device no larger than a typical transistor.

Disadvantages

Although semiconductor components have many advantages over the vacuum tubes that were once widely used, they also have several inherent disadvantages. First, semiconductor components are highly susceptible to changes in temperature and can be damaged if they are operated at extremely high temperatures or beyond their power dissipation specifications. Additional components are often required simply for the purpose of stabilizing circuits utilizing semiconductors so that they will operate over a wide temperature range. Semiconductor components may be easily damaged by exceeding their power dissipation limits and they may also be damaged when their normal operating voltages are reversed.

There are still a few areas where semiconductor devices cannot replace tubes. This is particularly true in high power, radio frequency applications. Nearly every commercial radio station has at least one high power vacuum tube in the final amplification stage of their transmitter. However, as semiconductor technology develops, these limitations will be overcome.

Despite the few disadvantages, semiconductor components are still the most efficient and reliable devices to be found. They are used in all new equipment designs and new applications are constantly being found for these devices in the military, industrial, and consumer fields. The continued use of semiconductor materials to construct new and better semiconductor components is almost assured because the techniques used are constantly being refined thus making it possible to obtain even superior components at less cost.

Semiconductors have had a profound effect on the design and application of electronic equipment. Not only have they greatly improved existing equipment and techniques by making them better and cheaper, but also they have permitted us to do things that were not previously possible. Semiconductors have revolutionized the electronic industry and they continue to show their greater potential. Your work in electronics will always involve semiconductor devices. Examine this unit on semiconductor fundamentals carefully and you will benefit from the resulting knowledge.

CONDUCTION IN INTRINSIC GERMANIUM AND SILICON

Although rarely used in their pure form, it is important to have a basic understanding of conduction in a semiconductor. It is beyond the scope of this primer to delve into a detailed discussion of conduction any modern materials science, solid state physics, or inorganic chemistry textbook will develop the theory explicitly.

Conduction of electrons in a material is governed by the energy difference between a band of orbitals known as the valence band and another band of orbitals called the conduction band. The theory behind the source of the valence and conduction bands are rooted in molecular orbital theory and quantum electrodynamics. An electron in a substance can only move while in the conduction band in the presence of an electric field.

A conductor, such as a metal, is a good conductor of electricity because the valence band of orbitals overlaps the conduction band. Electrons can be easily promoted to the conduction band and transported along the wire. In an insulator, such as glass, the conduction band is too far away in energy from the valence for electrons to be easily promoted into the conduction band. In an insulator, the conduction band may be separated so far from the valence band (in energy) that a spark will be produced preferentially rather than the electron enter a conduction band.

In a semiconductor, on the other hand, the conduction band is separated from the valence band by only a small amount of energy. An electron can be stimulated to enter the conduction band using an applied potential or even light or heat. While in the conduction band, the electron will move as in any conductor. However, if the applied potential falls below the necessary level to cause the electron to promote to the conduction band, the semiconductor material will act more as an insulator.

Low Temperature Characteristics

At low temperatures, the valence electrons are held tightly to their parent atoms and are not allowed to drift through the crystalline structures of the semiconductor material. Since the valence electrons are not free to drift from one atom to the next, the material cannot support current flow. At extremely low temperatures, pure germanium and silicon crystals function as insulators.

High Temperature Characteristics

As the temperature of a germanium or silicon crystal is increased, the valence electrons within the material become agitated and some of them will occasionally enter the conduction band. A small number of electrons will be free to drift from one atom to the next in a random manner. These free-moving electrons or *free electrons* are able to support a small amount of electrical current if a voltage is applied to the semiconductor material. In other words, as the temperature of the semiconductor material increases, the material begins to acquire the characteristics of a conductor. For all practical purposes, however, enough heat energy is available even at room temperature to produce a small number of free electrons that can support a small amount of current. Only when the semiconductor materials are exposed to extremely high temperatures, can a point be reached where they will conduct current as well as an ordinary conductor. Under normal conditions, however, this high temperature usage of semiconductors is never encountered.

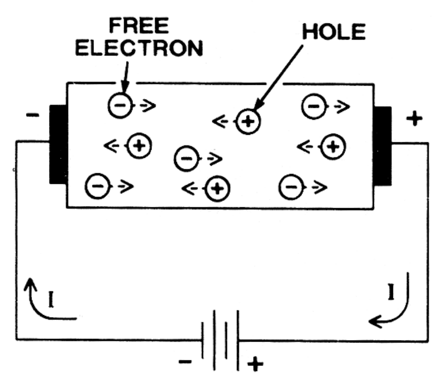
Holes

To understand exactly why a semiconductor is able to allow current to flow, we must take a closer look at the internal structure of the material. When an electron leaves the valence band an open space or vacancy is left behind. The space that was previously occupied by the electron is generally referred to as a *hole.* A hole simply represents the absence of an electron. Since an electron has a negative charge, the hole represents the absence of a negative charge. This means that the hole has the characteristics of a positively charged particle. Each time an electron leaves the valence band, a positive hole is created. Each corresponding electron and hole is referred to as an electron-hole pair.

The number of electron-hole pairs produced within a semiconductor material will increase with temperature. However, even at room temperature, a small number of electron-hole pairs will exist. Some of the free electrons will tend to drift randomly and the holes will try to absorb some of the electrons. This means that some electrons will simply jump from hole to another. If an electron jumps from one valence orbital to fill in a hole, another hole is created when the electron leaves the orbital. The hole, therefore, *appears* to move in the opposite direction of the electron. If another electron moves into the hole that was just created, another hole is produced and the previous hole appears to move randomly through a pure semiconductor material.

Current Flow

When a pure semiconductor material such as germanium or silicon is subjected to an electric potential, the negatively charged free electrons are attracted to the positive terminal of the voltage source. The positive holes that are created by the free electrons *appear* to drift toward the negative terminal of the voltage source. As the free electrons flow into the positive terminal of the voltage source, an equal number of electron leave the negative terminal of the voltage source. These electrons are injected into the semiconductor material where many of these electrons are captured or absorbed by holes. As the holes and electrons recombine in this manner, the holes cease to exist. Therefore the holes constantly drift to the left and then disappear while the electrons flow to the right side of the material where they are then drawn out of the material and into the positive terminal of the voltage source.



It is important to remember that current flow in a semiconductor material consists strictly of only electrons moving in the electric field (no different than a wire); however, it appears that the current is due to both electrons and holes. The holes appear like positively charged particles while the electrons are actually negatively charged particles. The holes and electrons flow in opposite directions and the number of electron-hole pairs produced within a material increase as the temperature of the material increases. Since the amount of current flowing in a semiconductor is determined by the number of electron-hole pairs in the material, the ability of a semiconductor material to pass current increases as the temperature of the material increases.

The temperature-current relationship makes a semiconductor appear to be a resistor with a negative thermal coefficient of resistivity; that is, the resistance decreases with increasing temperature. This observation has a significant ramification in circuits which utilize semiconductor devices. If the semiconductor device is allowed to pass sufficient current such that it heats up, the resistance of the device will decrease thus allowing more current to flow. Since power dissipation increases as the square of current but only linearly with resistance, the ultimate effect is that of *thermal runaway*. The increasing current causes the semiconductor device to dissipate even more heat which further lowers the resistance of the device. This cyclical effect will often result in the device quickly exceeding its power or current rating with the commensurate destruction of the device. In most circuits utilizing semiconductor devices such as transistors and integrated circuits, some sort of thermal protection is usually designed into the circuit.

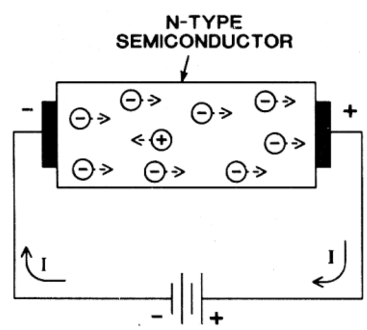
CONDUCTION IN DOPED GERMANIUM AND SILICON

Pure semiconductor materials contain only a small number of electrons and holes at room temperature and therefore conduct very little current. However, the conductivity of these materials can be increased considerably by a process known as *doping*. Pure semiconductor materials such as germanium and silicon are doped by adding other materials to them when they are produced. Basically there are two types of dopants that are added to germanium and silicon crystals: 1) a pentavalent material (such as arsenic) which has five valence electrons and 2) a trivalent material (such as gallium) which has three valence electrons. Doping may only be in the parts-per-billion (ppb) or parts-per-million (ppm) range but will have a profound effect on the properties of the semiconductor.

n-Type Semiconductors

When a pure semiconductor material is doped with a pentavalent element, such as arsenic (As), some of the atoms in the crystal lattice structure of the tetravalent semiconductor are replaced by arsenic atoms. The arsenic atom replaces one of the semiconductor atoms and is sharing four of its valence electrons with adjacent atoms in the crystal lattice. However, the fifth electron is not part of a bond and can be easily freed from the atom. This arsenic atom is called a donor atom because it donates a free electron to the crystal lattice. Even though there are donor atoms in only the ppb or ppm range, there are many free electrons in the semiconductor.

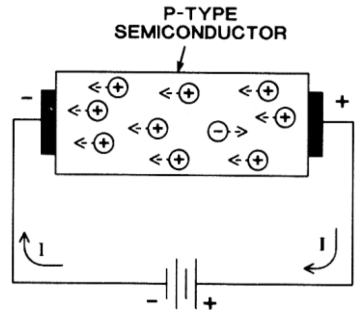
If a voltage is applied to an *n*-type semiconductor, the free electrons contributed by the donor atoms will flow toward the positive terminal of the battery. However, some additional free electrons will also flow toward the positive terminal. These additional free electrons are produced as electrons move to the conduction band and create electron-hole pairs and is identical to the action which takes place in a pure semiconductor material. The corresponding holes which are produced are then moved toward the negative terminal. Notice in the figure to the right, which summarizes conduction in an *n*-type semiconductor, more electrons than holes are present.



At normal room temperature the number of free electrons provided by the donor atoms will greatly exceed the number of holes and electrons that are produced by the breaking of covalent bonds. This means that the number of electrons flowing in the n-type semiconductor will greatly exceed the number of holes. The electrons, being in the majority, are therefore referred to as the majority carriers while the holes, which are in the minority, are referred to as the minority carriers.

p-Type Semiconductors

Doping pure semiconductor material with a trivalent element, such as gallium (Ga), causes some of the semiconductor tetravalent atoms to be displaced by trivalent atoms. This results in a hole in the crystal lattice. A large number of holes are present in the semiconductor because many trivalent atoms have been added. These holes readily accept electrons from other atoms. However, when a given hole is filled by an electron from another atom, the electron leaves another hole. Therefore, the holes drift through the lattice in the direction opposite to that of electron movement. The acceptor atoms remain fixed within the crystal lattice even though the holes can move freely.



If a voltage is applied to a *p*-type semiconductor as shown in the figure on the right, the holes provided by the acceptor atoms move from the positive to the negative terminal. These holes move in the same manner previously described. In other words as each electron moves into a hole, another hole is created in its place and since the electrons are attracted to the positive terminal, the holes move in the opposite direction.

In addition to the holes provided by the acceptor atoms, many additional holes are also produced in the *p*-type semiconductor material as electrons enter the conduction band thereby creating electron-hole pairs.

Under normal conditions, the number of holes provided by the acceptor atoms will greatly exceed the number of holes and electrons that are produced by placing a voltage across an *p*-type semiconductor. The number of holes "flowing" in the *p*-type semiconductor will therefore, greatly exceed the free electrons in the material. The holes being in the majority are referred to as the majority carriers and the electrons, which are in the minority, are referred to as the minority carriers.

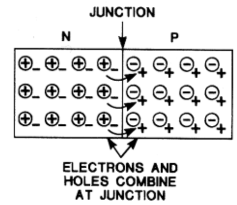
It is important to note that the semiconductor materials just described are referred to as *n*-type or *p*-type semiconductors on the basis of the identity of the majority carriers within these materials. The *n*-type and *p*-type materials themselves are not charged. In fact, both materials are electrically neutral.

It is also important to understand that *n*-type and *p*-type semiconductors have a much higher conductivity than pure semiconductors. Also, the conductivity of these materials can be increased or decreased by simply adding more or less dopant during the manufacturing process. The more heavily a semiconductor is doped, the lower its electrical resistance.

THE *p-n* JUNCTION: DIODES

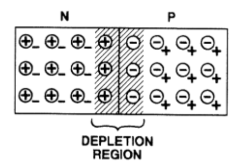
Junction Diodes

Let's consider the action which takes place when doped semiconductors are combined to form a diode. Basically, a diode is created by joining *n*-type and *p*-type semiconductors as shown on the right. When these oppositely doped materials come in contact with each other, a junction is formed where they meet. Such diodes are referred to as *junction diodes* and can be made of either silicon or germanium. When the junction is formed, a unique action takes place.



Depletion Region

The mobile charges, free electrons and holes, in the vicinity of the junction are strongly attracted to each other and therefore drift toward the junction. Some of the free electrons move across the junction and fill some of the holes in the *p*-type material. As the free electrons cross the junction, the *n*-type material becomes depleted of electrons in the vicinity of the junction. At the same time, the holes within the *p*-type material become filled and no longer exist. This means that the *p*-type material also becomes depleted of holes near the junction. This region near the junction where the electrons and holes become depleted is referred to as the *depletion region.* The depletion region extends for only a short distance on each side of the junction as shown in the figure.



Recall that electrons are the majority carriers in the *n*-type material and holes are the majority carriers in *p*-type materials. Therefore, no majority carriers exist within the depletion region. Also, it is important to note that the *n*-type and *p*-type materials are no longer neutral or uncharged. In other words the *n*-type material has lost electrons thus causing the positive donor atoms to outnumber the free electrons. The *n*-type material therefore takes on a positive charge near the junction. The *p*-type material has lost holes which means that the negatively charged acceptor atoms will outnumber the holes. The *p*-type material therefore, takes on a negative charge near the junction. The result is that opposite charges now exist on each side of the junction.

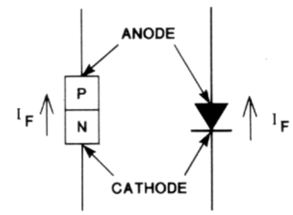
The depletion region does not continue to become larger and larger until the *n*-type and *p*-type materials are completely depleted of majority carriers. Instead, the action of the electrons and holes combining at the junction tapers off very quickly. Therefore, the depletion area remains relatively small. The size of the depletion region is limited by the opposite charges which build up on each side of the junction. The negative charge which accumulates in the *p*-type material eventually becomes great enough to repel the free electrons and prevent them from crossing the junction. The positive charge which accumulates in the *n*-type material also helps to stop the flow of free electrons by attracting and holding them back so they cannot move across the junction.

Barrier Voltage

The opposite charges that build up on each side of the junction create a difference in potential or voltage. This difference in potential effectively limits the size of the depletion region by preventing the further combination of electrons and holes. It is referred to as the *potential barrier* or the *barrier voltage*.

The barrier voltage produced within a *p-n* junction will usually be on the order of several tenths of a volt. For example, a *p-n* junction made from doped silicon will have a typical barrier voltage of ~0.6 volts while a *p-n* junction made from doped germanium will have a typical barrier voltage of ~0.3 volts. Although the barrier voltage exists inside of the junction diode and therefore cannot be measured directly, its presence becomes apparent when an external voltage is applied to the diode. Later in this unit you will see why this is true.

Diode Symbols

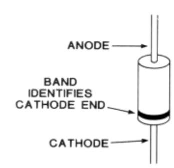


The schematic symbol most commonly used to represent the diode is shown on the right along with the equivalent representation of *p-n* junction diode. Notice that the *p* section of the diode (anode) is represented by an arrow and the *n* section (cathode) is represented by a bar. The arrows that are placed beside the diode and its symbol indicate the direction of forward current (I_F) or electron flow. However, they are now commonly used to describe the two sections of a junction diode. The cathode (*n*-type) is simply the section of the diode that supplies the electrons and the, anode (*p*-type) is the section that collects the electrons.

Investigation 18: Characterizing a Silicon Diode with an Ohmmeter

DMM
1N4001 silicon diode (any silicon diode is acceptable)

Procedure



1. Examine the diode. It should have a band nearer one end than the other, as in the diagram above. The band signifies the cathode end of the diode.

2. With the DMM in ohmmeter mode and set to the lowest resistance range setting, touch the red lead (+) to the anode and the black lead (-) to the cathode[12]. If necessary, adjust the ohmmeter range until the display reads a resistance. What is the resistance of the diode?

3. Reverse the leads on the diode and set the ohmmeter to the highest range setting. When

[12] Recall that the ohmmeter uses an internal battery to provide current. The color of the leads shows which pole of the battery to which they are connected.

the black lead is on the anode and red lead on the cathode, does the diode have high or low resistance?

Discussion

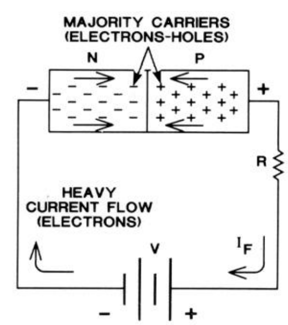
Since the resistance is lowest when the electron source (negative lead) is attached to the cathode, you have discovered that the forward current must flow from the *n*-junction (cathode) to the *p*-junction (anode) of the diode. This means that the forward electron current through the symbol must flow against the arrow. For a diode to conduct, the voltage applied to the diode must be the correct polarity and amplitude. The voltage and its polarity is referred to the *bias voltage*.

DIODE BIASING

Whenever diodes are used in electronic circuits, they are subjected to various voltages and currents. The polarities and amplitudes of the voltages and currents (the *bias*) must be such that proper diode action takes place.

Forward Bias

In the previous discussion, it was shown that free electrons and holes, the majority carriers, combine at the *p-n* junction to produce a depletion region. The depletion region represents an area that is void of majority carriers but at the same time contains a number of positively charged donor atoms and negatively charged acceptor atoms. The positive and negative charges are separated at the junction and effectively create a barrier voltage which opposes any further combining of majority carriers. Keep in mind that this action takes place in a *p-n* junction diode that is not subjected to any external voltage.



When a *p-n* junction diode is subjected to a sufficiently high external voltage as shown in the figure, the device will function in a somewhat different manner. Notice that the negative terminal of the voltage source is connected to the *n*-junction of the diode. An external resistor, R, is used to limit the current. Under these conditions the free electrons in the *n*-junction are repelled by the negative battery terminal and forced toward the *p-n* junction where they will neutralize the positively charged donor atoms in the depletion region. During this same period of time, the free electrons that had initially accumulated to create a negative charge on the *p* side of the junction are attracted toward the positive battery terminal. Therefore, the negative charge on the *p* side of the junction is also neutralized. This means that the positive and negative charges which form the internal barrier voltage are effectively neutralized and no barrier voltage will be present to stop the combining of majority carriers at the junction. The *p-n* junction diode is therefore able to support a continuous flow of current at this time. This action will occur only if the supply voltage exceeds the barrier potential (~0.6 V for a silicon junction diode).

Since the diode is now subjected to an external voltage, a constant supply of electrons flow into the *n*-junction of the diode. These electrons drift through the *n*-type material toward the junction.

The movement of these electrons through the *n*-junction is sustained by the free electrons (majority carriers) that exist within this material. At the same time, the holes (majority carriers) in the *p*-junction also drift toward the *p-n* junction. The electrons and holes combine at the *p-n* junction and effectively disappear as they neutralize each other. However, as these electrons and holes combine and are effectively eliminated as charge carriers, new electrons and holes appear at the outer edges of the *n*- and *p*-junctions. The majority carriers therefore continue to move toward the *p-n* junction as long as the external voltage is applied.

It is convenient to analyze the action which takes place in the *p*-junction of the diode by considering the movement of holes instead of electrons. However, it is important to realize that electrons do actually flow through the *p*-type material. The electrons are attracted by the positive terminal of the battery and as the electrons leave the *p*-junction and enter the power supply, holes are created at the outer edge of the *p*-junction. These holes drift toward the *p-n* junction where they combine with electrons and effectively disappear.

When the diode is in the conducting state (*i.e.* positive voltage on the anode), it is said to be conducting current in the *forward* direction and is considered to be *forward-biased* by the external voltage. The current which flows through the forward-biased diode is limited by the resistance of the *p*-type and *n*-type semiconductor materials as well as any external resistance. Normally the diode resistance is quite low. Connecting a forward bias voltage directly to the diode may result in a current large enough to generate sufficient heat to destroy the diode. For this reason, forward-biased diodes are usually connected in series with a resistor or other current limiting device.

A forward-biased diode will conduct current as long as the external bias voltage is sufficiently high and the polarity is correct. For example, a forward biased germanium diode requires approximately 0.3 V to drive the diode into conduction. Silicon diodes require a forward bias of approximately 0.7 volts in order to begin conducting. The external voltage applied to the diode must be large enough to neutralize the depletion area and therefore neutralize the barrier voltage that exists across the *p-n* junction of the diode. Once this internal voltage is overcome, the diode will conduct in the forward direction.

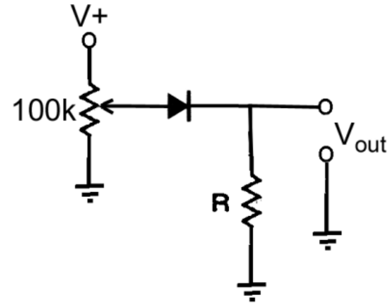
Once the diode is conducting, a voltage will be dropped across the device. This occurs because the diode's semiconductor material has a low but finite resistance value and the current flowing through it must produce a corresponding voltage drop. As it turns out, this forward bias voltage drop is approximately equal to the barrier potential. This is 0.3 volts for a germanium diode and 0.7 volts for a silicon diode.

Investigation 19: Characterizing the Current-Voltage Relationship of a Forward Biased Silicon Diode

Trainer
DMM (2 DMMs will make measurements more convenient)
1N4001 silicon diode (any silicon diode is acceptable)
1kΩ resistor
100 kΩ potentiometer (you may use the potentiometer on the trainer)

Procedure

Record your data.



1. Set up the circuit shown. Set the positive DC supply on the Trainer to full output voltage and use the 100 kΩ potentiometer to adjust the voltage to the diode resistor network. The regulated power supply in the Trainer may not be able to be set to 0 V but, using the potentiometer to control voltage, it will be possible to get nearly 0 V applied to the diode-resistor network.

2. Adjust the 100 kΩ potentiometer to provide 0 V (vs. ground) to the diode-resistor network. Measure and record this voltage. Break the circuit between the resistor and diode and measure the current through the network. If you have 2 DMMs, leave one in ammeter mode and in circuit. Use the other DMM to measure voltage drops. Record the applied voltage, voltage drop across the diode, and the network current.

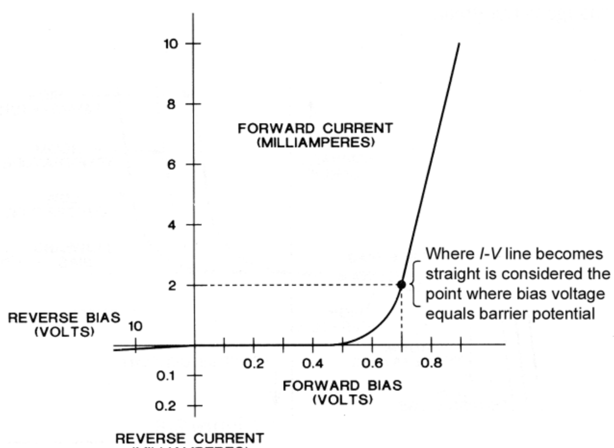
3. Increase the applied voltage to about 0.2 V (vs. ground). Measure and record the applied voltage, the voltage drop across the diode, and the network current. Repeat the measurements, increasing the applied voltage by about 0.2 V each time, to a final applied voltage of about 5 V.

4. Reverse the diode (but leave the leave the ammeter and voltmeter leads in the same position as in step 3) and repeat measurements in step 3 with 1 V increments.

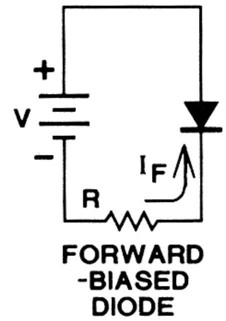
Make a graph of network current vs. bias voltage (applied voltage). At approximately what voltage does the diode start conducting? What is the difference in currents in the forward and reverse bias voltage condition?

Discussion

You might have obtained a graph which looks similar that shown below. (Notice the change in scale of the reverse bias axes.)



The amount of forward bias current, I_F, is a function of the applied DC bias voltage, V, the forward voltage drop, V_F, and the external resistance, R. The relationship simply involves Ohm's law



$$I_F = \frac{V - V_F}{R}$$

For example, the forward current in a silicon diode (barrier potential approximately 0.7 V)with a bias voltage of 10 V and an external resistor of 100 ohms is

$$I_F = \frac{10\ V - 0.7\ V}{100\ \Omega} = 0.093\ A$$

Reverse Bias

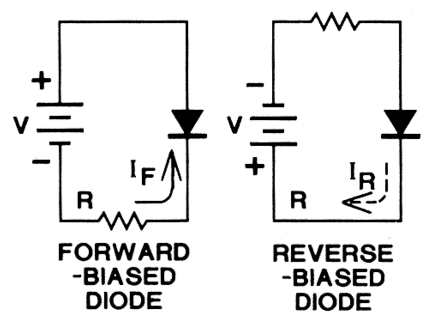
As you discovered in the first part of Investigation 19, a forward-biased diode is able to conduct current in the forward direction because the external bias voltage forces the majority carriers together so that they can combine at the junction of the diode and create a continuous flow of current. In order to achieve this condition the negative terminal of the battery is connected to the *n* section of the diode and the positive terminal is connected to the *p* section. However, if the source connections are reversed (as in the latter part of Investigation 19) the diode will operate in a different manner. The negative terminal of the battery is now connected to the *p* section of the diode while the positive terminal is connected to the *n* section. The diode is now considered to be *reverse-biased.* Under these conditions the free electrons in the *n* section will be attracted toward the positive battery terminal thus leaving a relatively large number of positively charged donor atoms in the vicinity of the *p-n* junction. In fact, the number of positive ions in the *n*-junction at times will even outnumber the positive ions that exist in an unbiased diode. This effectively increases the width of the depletion region on the *n* side of the *p-n* junction thus causing the positive charge on this side of the *p-n* junction to increase. At the same time, a number of electrons leave the negative terminal of the battery and enter the *p* section of the diode. These electrons fill the holes near the junction thus causing the holes to move toward the negative terminal. A large number of negatively charged acceptor atoms are therefore created near the *p-n* junction. This effectively increases the width of the depletion region on the *p* side of the *p-n* junction. These opposite charges will build up until the internal barrier voltage is equal and opposite to the external battery voltage. Under these conditions the holes and electrons (majority carriers) cannot support current flow and the diode effectively stops conducting.

Actually, an extremely small current will flow through the reversed-biased diode. This small current is sometimes referred to as a *leakage current* or a *reverse current* and is designated as I_R. Recall that the minority carriers are holes in the *n*-type material and electrons in the *p*-type material. When the diode is reverse-biased, the minority carriers are forced toward the junction where they combine and thus support an extremely small current. This action closely resembles the action which takes place in the forward-biased diode but it is on a much smaller scale. Leakage currents are only microamperes in germanium diodes and nanoamperes in silicon diodes. These currents are usually many orders of magnitude less than the usual forward current. Note that germanium diodes normally produce a higher leakage current than silicon diodes, a disadvantage of germanium diodes often offset by a lower barrier potential and forward voltage drop.

The number of minority carriers in the *n* and *p* materials is extremely small at room temperature. However, as temperature increases, a greater number of electron-hole pairs are generated within the two materials. This causes an increase in minority carriers and a corresponding increase in leakage current.

Summary of Biasing

The figure shows how forward-biased and reversed-biased diodes are represented in schematic form. Notice that when the negative and positive terminals of the battery are connected to the cathode and anode of the diode, respectively, the diode is forward-biased and will conduct a relatively high forward current, I_F. The resistor is added in series with the diode as shown to limit this forward current to a safe value. Also notice that when the battery terminals are reversed, the diode is reverse-biased and only a very low reverse current, I_R, will flow through the device.

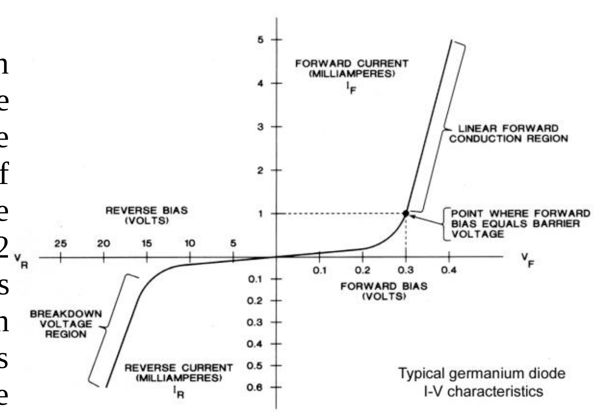


The operation of a *p-n* junction diode can be summed up in this manner. The diode is a unidirectional electrical device since it conducts current in only one direction. When it is forward biased, current flows through it freely since it acts as a very low resistance. When the diode is reversed biased, current does not flow through it. It simply acts as an open circuit or extremely high resistance. Only a small, temperature sensitive leakage current flows in the reverse biased condition. The diode is effectively a polarity sensitive electrical switch. When forward biased the diode switch is closed. When reverse biased it is open.

Germanium Diode

Forward characteristics

The I-V curve for the germanium diode is similar in appearance to the curve obtained for the silicon diode in Investigation 19. Until the forward bias voltage across a germanium diode increases beyond a value of 0.2 V, the forward current through a germanium diode is extremely small and almost insignificant. Beyond 0.2 V, the forward current increases as the forward bias voltage is increased still further. The increase in forward current really starts to occur as the diode goes into conduction when the external bias voltage



Typical germanium diode I-V characteristics overcomes the diode's internal barrier voltage. Once the bias voltage exceeds the barrier voltage (0.3 volts), the forward current increases very rapidly and at a linear rate because the diode then acts as a low resistance. Throughout the linear portion of the curve, the voltage across the diode is only several tenths and, while the forward voltage drop is not constant, it changes very little over a wide current range.

Reverse Characteristics

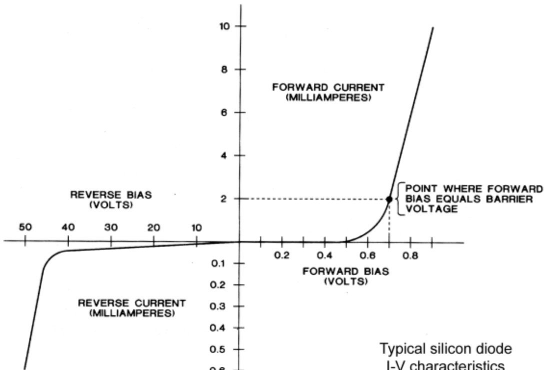
The I-V curve for a typical germanium diode also shows that when the diode is reverse-biased, the reverse current that flows is extremely small (but larger than a silicon diode). Notice that the reverse current increases slightly as the reverse voltage increases but remains less than 0.1 mA until the reverse voltage approaches a value of 20 volts. Then the reverse current suddenly increases to a much higher value. This sudden increase in reverse current results because the reverse bias voltage becomes strong enough to tear many valence electrons from their parent atoms and therefore increase the number of electron-hole pairs in the *n* and *p* materials. This causes an increase in minority carriers which in turn support a higher reverse current. In other words the junction simply breaks down when the reverse bias voltage approaches a value of 20 volts.

The voltage at which the sudden change occurs is commonly referred to as the *breakdown voltage.* This breakdown voltage will vary from one diode to the next since it is determined by the exact manner in which the diode is constructed. In most cases, ordinary germanium diodes are permanently damaged when breakdown occurs; however, there are special diodes which are designed to operate in this region. These special devices, known as zener diodes. When breakdown occurs, the diode no longer offers a high resistance to the flow of reverse current and therefore cannot effectively block current in the reverse direction. For these reasons, operation in the breakdown region is avoided when an ordinary *p-n* junction diode is being used.

Silicon Diode

While a silicon diode operates the same as the germanium diode, there are some important differences in their characteristic curves. Let's study some these differences in a little detail.

Forward Characteristics



Typical silicon diode I-V characteristics

The I-V curve shows the characteristics of a typical silicon diode. Notice that the forward characteristics of this diode are basically similar to those of the germanium diode; however, there is an important exception. The internal barrier voltage of the silicon diode is not overcome until the forward bias voltage is equal to approximately 0.7 V. Beyond this point the forward current increases rapidly and linearly. The corresponding forward voltage across the diode increases only slightly. The exact amount of forward voltage required to overcome the barrier voltage will vary from one silicon diode to the next but will usually be close to the 0.7 V.

Reverse Characteristics

The reverse characteristics of the silicon diode are also similar to those of the typical germanium diode. However, the silicon diode has a much lower reverse current than the germanium type. Observe that the reverse current remains well below 0.1 milliamperes until the breakdown voltage of the device is reached. Then, as with the germanium unit, a relatively high reverse current is allowed to flow. A breakdown voltage of 45 volts is indicated in the figure, however, this voltage will vary from one silicon diode to the next. Also, the reverse currents in many silicon diodes may be in the extremely low nanoampere range and therefore insignificant for most practical applications. Are any of these observations consistent with the results you obtained in Investigation 19?

Diode Ratings

When the important characteristics of silicon and germanium diodes are compared it becomes apparent that either type can be damaged by excessive forward current. For this reason, manufacturers of these diodes usually specify the maximum forward current (I_{max}) that each type can safely handle. Also, both types can be damaged by excessive reverse voltages which cause the diode to breakdown and conduct a relatively large reverse current. To insure that the various diodes are not subjected to dangerously high voltages, manufacturers of these devices usually specify the maximum safe reverse voltage that can be applied to each particular device. This maximum reverse voltage is commonly referred to as the *peak inverse voltage,* which is usually abbreviated PIV.

Temperature Considerations

In some critical applications it is also necessary to consider the effect that temperature has on diode operation. In general, the diode characteristic that is most adversely affected by changes in temperature is the diode's reverse current. Notably, the germanium diode is affected more than the silicon diode. At extremely low temperatures the reverse current through a typical diode will be practically zero. But at room temperature this current will be somewhat higher although still quite small. At extremely high temperatures an even higher reverse current will flow which in some cases might interfere with normal diode operation.

The forward voltage drop across a conducting diode is also affected by temperature changes. The forward voltage drop is inversely proportional to temperature. As the temperature rises, the voltage drop decreases. This effect is the same in both germanium and silicon devices.

DIODE APPLICATION: RECTIFICATION

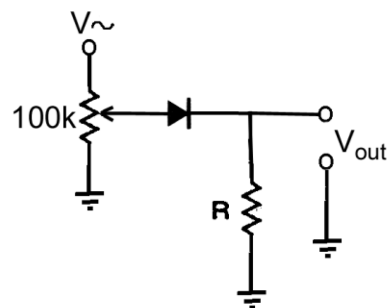
Junction diodes are used extensively for their property of conduction in only one direction. Many applications exist for just this property but one application stands out – *rectification*. Rectification is the process of converting alternating current into direct current. Diodes are uniquely suited to perform this task and, as such, are often called rectifiers.

Most electronic circuits require DC for operation but are powered from an AC line. Rectification of the AC is therefore necessary to convert the available power (usually 110 V in the United States) into DC which may be as low as a few volts or as high as thousands of volts. Normally, a rectifier circuit will not use a diode alone in the rectification process but will incorporate additional components such as transformers (for stepping up or stepping down the primary voltage), resistors (for limiting current), and capacitors (for filtering the rectified voltage).

Investigation 20: The Half-Wave Rectifier

Trainer
DMM
Oscilloscope (with two 10x probes)
1N4001 silicon diode (any silicon diode is acceptable)
1kΩ resistor
100 kΩ potentiometer (you may use the potentiometer on the trainer)

Procedure



Record your data.

1. Set up the circuit shown. Use the 100 kΩ potentiometer to adjust the 15 VAC from the Trainer. Pay particular attention to the polarity of the diode.

2. Adjust the output voltage of the potentiometer to about 5 V_{rms}.

3. Attach channel 1 of the oscilloscope to the output of the potentiometer and channel 2 to the point labeled V_{out}. Be sure to DC couple both inputs.

4. Analyze the waveforms. Measure the V_p of the source waveform (it may be necessary to detach one of the diode's leads to get a correct waveform). Measure the V_p of the output waveform.

5. Reverse the diode and repeat the characterization of the circuit.

6. Do not disassemble this circuit yet.

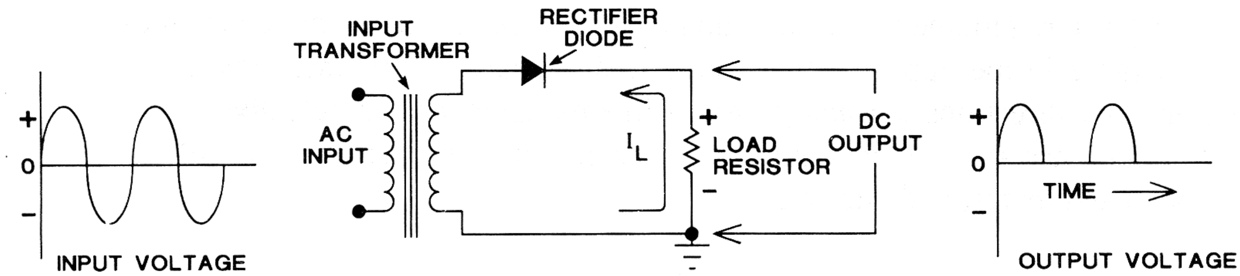
Explain the shape of the output waveforms. Why is only one-half of the input waveform observed. Is the output voltage lower than the input voltage? Why?

Discussion

The circuit studied is known as a half wave rectifier since it conducts during only one-half of the

AC input.

Examine the following circuit which uses a step-down transformer to convert 110 VAC to 5 VAC and a silicon diode. This circuit illustrates the basic half-wave rectifier. The resistor shown is the circuit load (and is represented by the 1 kΩ resistor in Investigation 20.



Since the diode is in series with the load, current will flow in only one direction; that is, only when the diode is forward biased during each alternation. The pulsating DC output voltage produced is positive with respect to ground because current can only flow when the anode of the diode is positive with respect to ground. By reversing the diode, the sign of the output voltage is reversed since the forward bias voltage is now reversed with respect to ground.

This circuit provides the basis for constructing AC to DC power conversion. However, this simple circuit is usually not adequate to provide power to most circuits which require DC. Most DC powered circuits require pure or continuous DC for proper operation. Thus, additional components are necessary to further refine and polish the output signal.

The pulsating DC produced by the circuit shown and studied in Investigation 20 can be approximated as DC with low frequency noise (60 Hz). It should be possible, then, to remove the pulsation and produce nearly pure DC by using a low-pass filter (with <60 Hz cutoff) prior to the load. The simplest low frequency low-pass filter is a large capacitor connected from V_{out} to ground.

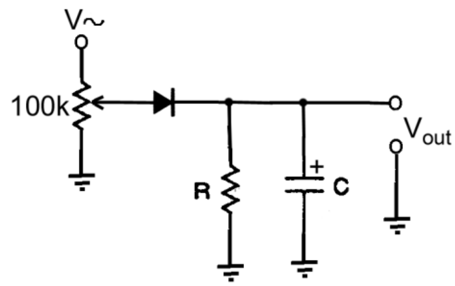
Investigation 21: The Half-Wave Rectifier with Filtering

Trainer
DMM
Oscilloscope (with two 10x probes)
1N4001 silicon diode (any silicon diode is acceptable)
1kΩ resistor
100 – 2000 μF (25 V or greater) electrolytic capacitor
100 kΩ potentiometer (you may use the potentiometer on the trainer)

Procedure

Record your data. This circuit is similar to the one used in Investigation 20.

1. Set up the circuit shown. Use the 100 kΩ potentiometer to adjust the 15 VAC from the Trainer. Pay particular attention to the polarity of the diode and the capacitor.



2. Adjust the output voltage of the potentiometer to about 5 V_{rms}.

3. Attach channel 1 of the oscilloscope to the output of the potentiometer and channel 2 to the point labeled V_{out}. Be sure to DC couple both inputs.

4. Observe and record the output waveform.

Is the output of the circuit with the capacitor different than the output of the circuit without the capacitor? What DC voltage is obtained (use the oscilloscope or DMM in DC mode to measure this)?

5. Set the oscilloscope to display only channel 2 (the DC output channel). Set the input coupling to AC (to remove the DC component). Adjust the input vertical sensitivity high enough to display the remaining residual AC "ripple" on the DC signal. Measure and record the ripple voltage.

Calculate the percentage ripple ($V_{ripple}/V_{DC} \times 100$) remaining on the DC signal.

Discussion

Understanding the role of the capacitor is simple if you recall the operation of the low-pass filter. The capacitor charges up to the peak value of the DC pulse during the time the diode is in conduction. During the next half cycle (when the diode is reversed biases and not conducting), the capacitor discharges into the load to help maintain a continuous current flow. The larger the capacitor, the lower the residual ripple remaining on the DC. With a sufficiently high capacitance, the DC voltage will be essentially constant.

OTHER DIODE APPLICATIONS

While rectification is the most prevalent use the common junction diode, other uses abound. Furthermore, there are a variety of different types of diodes, which will not be discussed here, that are in common use. Among these are the zener diode (a diode which is normally used in the reversed voltage bias direction at or beyond the breakdown voltage), the tunnel diode (which exhibits "negative" resistance in a certain region of its I-V curve), and the varactor (a diode with a voltage-variable capacitance), to name a few.

BIPOLAR JUNCTION TRANSISTORS