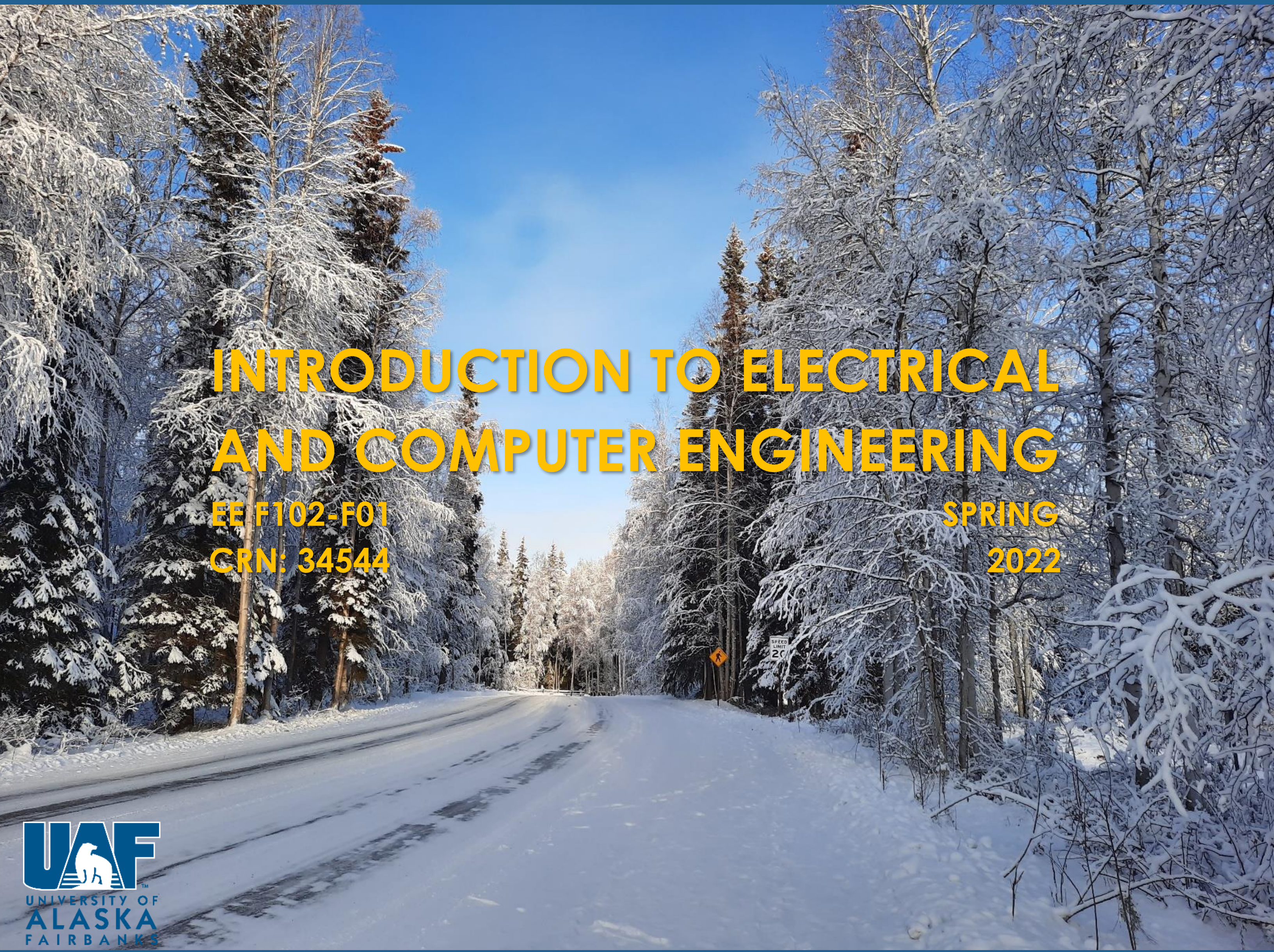
UNIVERSITY OF ALASKA FAIRBANKS

COLLEGE OF ENGINEERING AND MINES

Department of Electrical and Computer Engineering



INTRODUCTION TO ELECTRICAL AND COMPUTER ENGINEERING

EE F102-F01
CRN: 34544

SPRING
2022

UNIVERSITY OF ALASKA FAIRBANKS

Photo taken by Maher Al-Badri at the UAF campus on Sunday 01 NOV 2020 14:15
Temperature: -23 ºC (-9.4 ºF)

INTRODUCTION TO ELECTRICAL AND COMPUTER ENGINEERING

EE F102-F01 SPRING
CRN: 34544 2022

This text is prepared by

Denise Thorsen

Reviewed by

Maher Al-Badri

Definitions

Electrical Engineering
A field of engineering that deals with the study and application of electricity, electronics, and electromagnetism, covering a wide range of subfields including electronics, digital computers, power engineering, telecommunications, control systems, RF engineering, and signal processing (*source: Wikipedia*).

Computer Engineering
A discipline that integrates several fields of electrical engineering and computer science required to develop computer systems (*source: Wikipedia*).

Electrical Engineers
Electrical engineers design, build, and test electrical and electronic equipment, such as power generation systems, radar and navigation systems, communications systems, and computer systems, to name a few. Electrical engineers use their knowledge of electricity, electromagnetism, and electronics to develop creative solutions to meet a broad range of needs. Electrical systems touch every aspect of our daily lives: the alarm clock that wakes you up, the oven or microwave that cooks your meals, the cell phone that connects you to your friends, the automobile that gets you to school, the computer that helps you do your homework, the stereo system that helps you relax, and all the entertainment and gaming system that occupy your time. Electrical engineers make all that possible and more.

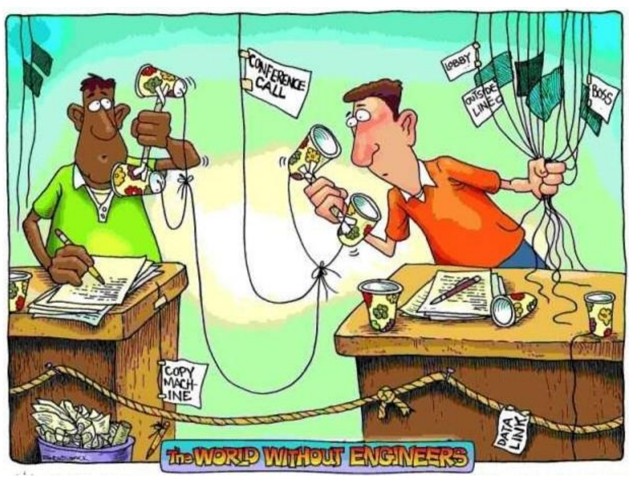


Figure 1. What life would be like without Electrical Engineers (EE)?!

Electrical engineers generally perform two parallel tasks: analysis of existing systems to diagnose failures and/or seek improvements; and design of new systems to meet specific electrical specifications.

- Analysis
  - Given a Circuit

    - Explain how the circuit works
    - Diagnose failures in the circuit

- Design
  - Given a problem description
    - Translate description into electrical specifications.
    - Convert electrical specifications to circuit.

Course Overview

EE102 *Introduction to Electrical and Computer Engineering* provides an overview of the electrical engineering sub-disciplines taught at UAF: power and control, communications, and computer engineering. The labs in EE102 are designed to model how practicing engineers might approach the design and development of a product. Through this course students develop most of the skills necessary to be successful in their future courses. The topics covered in this course are:

- **Basic Concepts**
  The course starts by defining the basic parameters used by electrical engineers: charge, current, voltage, power, energy, efficiency.

- **Circuit Analysis**
  The next three modules builds the knowledge tools used to analyze simple resistor circuits using independent and dependent sources. Students will learn more in depth circuit analysis including AC circuits, three-phase circuits, and network and system analysis in **EE203**.

- **Electric Machines**
  The following module jumps to magnetic circuits as a basis for learning about electric machines. In **EE303**, students will continue this topic adding electromechanical energy con- version principles, characteristics and applications of transformers, and synchronous and induction machines.

- **Electronic components**
  The next three modules introduce students to simple models of two common non-linear components, diodes and transistors and how those components may be used in circuits. More complicated models of these non-linear components are presented in **EE333** along with the basic principles of semiconductor devices and integrated circuits.

- **Digital Circuits**
  The course ends by designing diode and transistor circuits to develop digital logic gates and introduction to digital circuits. **EE343** provides more in depth principles and practices of digital design, including the analysis, design and implementation of combinational and sequential logic machines.

How to Get an A!

Instructional methods designed to promote conceptual understanding through interactive engagement of students in heads-on and hands-on activities have been shown [Hank, 1998] to increase course effectiveness well beyond that of traditional “lecture” methods. EE102 has been designed to promote interactive engagement of students through in-class problem-solving and discussions. **As such it is the student’s responsibility to come to class prepared to work by previewing the online material before class.** All students have the potential to get an ‘A’ in the class. It really depends on their level of effort. To be successful in this course you must:

- **Attend Class**
  Class time will be used to practice problems-solving and clarify areas of confusion.

- **Come to Class Prepared to Participate**
  Watch the lectures and read lecture materials posted on the Blackboard prior to coming to class. Develop a note sheet and bring it to class to help with the in-class problems. This is a participatory class; you get out of it what you put in.

- **Turn In Homework**
  There is a direct correlation between homework accomplished and final grade in class. Work in groups to complete the homework, but make sure you can answer all the homework problems by yourself. If you do not understand something **ASK QUESTIONS!**

- **Turn In Laboratory Reports**
  All lab reports are required for a passing grade. If you fail to turn in your lab reports you risk not passing the course regardless of your overall class standing.

- **Show All Work**
  When working problems on exams show all work. *The instructor grades primarily on process rather than final answer*, so if the instructor can’t follow your work, it is difficult to give you partial credit.

- **Identify Problems**
  If you have problems with the homework, lab, or class in general, talk to the instructor. Instructors do not read minds.

- **Classroom Conduct**
  Conduct yourself professionally in class and lab. Do your own work, that is the only way you will learn.
- **Participate! Practice! Practice! Practice!**

Frequently Asked Questions

- **Where do I find course information?**
  All course information, lectures, assignments, are accessed through the Course Blackboard.

- **When is the homework and/or lab report due?**
  The syllabus, which is posted on Blackboard, contains a course calendar which lists the due dates for all homework and lab reports.

- **When is the first/second/final exam?**
  The syllabus, which is posted on Blackboard, contains a course calendar which lists the dates for all exams.

- **How do I get help (homework/labs/general)?**
  The instructor is available during office hours and via email. Inform the instructor that you need to meet him/her. The instructor welcomes any student who would like to discuss any issue during the office hours. When students have a question, the instructor is there with an immediate answer. *However, due to the current COVID-19 pandemic, all meetings are to be conducted via zoom application. Zoom information can be found in the Course Outline document available on the the course Blackboard.* Also, the instructor is always available via email.

- **Why can't I pick who I work with?**
  One purpose of this class is to train you to work as a professional engineer. It is not often the case that you will have any choice in who you work with once you graduate. In this class students practice professionalism by working with a variety of students.

Historical Overview of Electrical Engineering

- 2600 BC Lodestone discovered.
- 600 BC Thales: Observes frictional electrification.
- 1269 Magnetic compass.
- 1600 Gilbert: uses term "electricity." Extensive work in electricity and magnetism and friction; earth = huge magnet.
- 1729 Gray: conductors-insulators.
- 1745 The Leyden Jar (storing electricity). franklin: electric origin of lightning. Positive and negative electricity.
- 1784 **Coulomb**: law of electric force. *Electricity is produced by friction or induction.*
- 1792 **Volta**: the battery; electricity produced from chemical erects.
- 1820 Aerstd - **Ampere**: discovery forces exerted on carrying conductors by magnetic fields. Biot - Savart: law.

- 1827 **Ohm**: relationship between voltage and current; resistance.
- 1831 **Faraday**: electromagnetic induction; transformer and generator. **Henry**: self-induction.
- 1847 **Kirchhoff's Laws**.
- 1858 First DC (Direct Current) generator.
- 1875 Crooke: experiments. Crooke's tube (→ oscilloscope, radiotubes, X-rays).
- 1880 Edison: incandescent lamp.
- 1886 First commercial transformer (AC; Alter- nating Current).
- 1888 Hertz: nature of electromagnetic waves.
- 1895 First AC installation converting mechani- cal energy (water) into electrical energy.
- 1904 Fleming: valve; diode detector.
- 1906 Lee DeForest: triode, amplifier. Thomson-Millikan: mass, charge of electron.
- 1945 First electronic computer built at MIT.
- 1947 Shockley-Bardeen-Brattain: transistor.
- 1958 Integrated circuits, microelectronics.
- 1971 First microprocessor (Intel).

Learning Objectives

At the end of this module you should:
- be able to successfully navigate the course.
- be able to determine the due dates for all assignments.
- understand what it takes to get the grade you want in the course.
- understand your responsibilities in the course.
- understand how to get help.

CANVAS Navigation

Announcements
Anything that happens during the semester that you need to know about will be announced here.

Course Syllabus
Everything you need to know about the course is located here: (1) the course syllabus, how to contact the instructor, how to contact the TA, office hours, semester calendar.

Lectures
The content of the course is contained in 24 Lectures. Lecture notes are provided for each lecture. You are required to read the lecture text prior to coming to class.

Homework

Homework will typically be assigned on a weekly basis. Homework information will be posted in this module where the date of assignment, submission due, and other information will be posted.

Quizzes

Quizzes will typically be delivered on a weekly basis. You will be informed about the materials you need to study for the quiz.

Laboratories

The first half of the semester you will progress through a series of connected Guided Labs which follow a simplified Engineering Design Cycle and lead to a culminating example design. During the second half of the semester you will be grouped in teams to formulate your own design following again the simplified Engineering Design Cycle. The Team Project will culminate in a presentation at the end of the semester.

Chapter 1

Review: Units, Numbers, and Waveforms

In this chapter you will learn about the SI system of units, both base units and derived units, and how to use their relationship to perform dimensional analysis. You will learn how to determine the accuracy and precision of numbers, sums and differences of numbers, and products and quotients of numbers in terms of their significant digits. You will learn the difference between scientific and engineering notation and the use of engineering prefixes. Finally, you will learn about conversion factors, how they are used and how they affect the accuracy and precision of the converted result.

1.1 System of Units

SI System (Systéme International) is the worldwide standard for systems of units which are used to describe, measure, and compare different physical quantities. The SI standard was established at the General Conference of Weights and Measures in Sévers, France in 1960. The United States, however, did not endorse the system until fifteen years later with the Metric Conversion Act of 1975.

1.1.1 Base Units

There are seven base units, shown in Table 1.1, that are regarded as dimensionally independent (meaning that they cannot be derived from other base units.) The first five are of particular interest for Electrical and Computer Engineers. We will be using these five throughout the semester.

Table 1.1: Table of base units.

Unit	**Abbreviation**	**Quantity**
meter	m	length
kilogram	kg	mass
seconds	s	time
ampere	A	electric current
kelvin	K	temperature
mole	mol	amount of a substance
candela	cd	luminosity

- **Meter-** Originally, the meter was defined as equal to one ten-millionth of the distance from the equator to the pole measured on a meridian. From 1889 to 1960, the meter was defined as the distance between two lines on a platinum-iridium bar preserved at the International Bureau of Weights and Measures near Paris. From 1960 to 1985, the meter was defined as 1,650,763.73 wavelengths of the orange-red radiation of krypton 86 under specific conditions. The meter is now defined as the distance traveled by light in a vacuum in $1/299,792,458$ seconds.

- **Kilogram** - The kilogram is currently the only SI unit that is still defined by an artifact rather than a fundamental physical property that can be reproduced in different laboratories. Originally, a kilogram was the mass of a cubic decimeter of water. In 1889 the international prototype of the kilogram, made of platinum-iridium, was created and is kept at the International Bureau of Weights and Measures. The kilogram is defined as a mass equal to the mass of the international prototype of the kilogram.

- **Second** - Early definitions of the second were based on the motion of the earth. 24 hours in a day meant that the second could be defined as $1/86,400$ of the average time required for the earth to complete one rotation about its axis. In the late 19^{th} century, astronomical observations revealed that this average time is lengthening, and thus the motion of the earth is no longer considered a suitable standard for definition. Since 1967, the second has been defined to be the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium 133 atom.

- **Ampere** - Quantitatively, the ampere is defined to be the constant current which will produce an attractive force of 2×10^{-7} newtons per meter of length between two straight, parallel conductors of infinite length and negligible circular cross section placed one meter apart in a vacuum. The definition is based on Ampere's force law which we will cover later in Chapter 7, Magnetism. The ampere is officially defined without reference to the quantity of electric charge.

- **Kelvin** - The kelvin is defined as 1/273.16 of the thermodynamic temperature of the triple point (equilibrium of solid, liquid, and gaseous phase) of pure water. In other words, it is defined such that the triple point of water is exactly 273.16 K. The zero point, absolute zero, on the Kelvin scale is the theoretical temperature at which the molecules of a substance have the lowest energy.

1.1.2 Derived Units

Other units can be derived from the seven base units by multiplying and dividing units within the system without numerical factors. Examples of derived units are shown in Table 1.2.

The unit of charge, the coulomb (C), is a derived unit. One coulomb is defined as the amount of electric charge carried past a point in one second by a current of one ampere. Conversely, a current of one ampere is one coulomb of charge going past a given point per second. Rather than defining the ampere in terms of the force between two current-carrying wires, it has been proposed to define the ampere in terms of the rate of flow of elementary charges, i.e. protons. Since a coulomb is approximately equal to $6.24150948 \times 10^{18}$ elementary charges, one ampere is approximately equivalent to $6.24150948 \times 10^{18}$ elementary charges moving past a boundary in one second. The proposed change would define 1 A as being the current in the direction of flow of a particular number of elementary charges per second. In 2005, the International Committee for Weights and Measures agreed to study the proposed change, and, depending on the outcome of experiments over the next

Table 1.2: Table of derived units.

Unit	**Abbreviation**	**Quantity**	**Base**
newton	N	force	$\frac{(\mathrm{kg \cdot m})}{\mathrm{s}^2}$
joule	J	work, energy, heat	$\mathrm{N \cdot m}$, $\frac{\mathrm{kg \cdot m^2}}{\mathrm{s}^2}$
watt	W	power	J/s
volt	V	electric potential, emf	J/C, W/A
coulomb	C	electric charge	$\mathrm{A \cdot s}$
ohm	Ω	resistance	V/A
farad	F	capacitance	C/V
henry	H	inductance	Wb/A
hertz	Hz	frequency	1/s
weber	Wb	magnetic flux	$\mathrm{V \cdot s}$
tesla	T	magnetic flux density	$\mathrm{Wb/m^2}$, $\frac{\mathrm{N}}{\mathrm{A \cdot m}}$

few years, to formally propose the change at the 25^{th} General Conference on Weights and Measures in 2014.

Some of the units shown are derived with respect to other derived units. For example, electric potential, or volt (V), is derived in terms of joules per coulomb or watts per ampere. Each of the units joule (energy) , coulomb (charge) and watt (power) are themselves derived units. We can determine the equation for the volt using solely base units by substituting in for joule, coulomb, and/or watt, to discover that the electric potential, volt, defined in terms of base units is:

$$V = \frac{J}{C} = (\frac{\mathrm{kg \cdot m^2}}{\mathrm{s^2}})/(\mathrm{A \cdot s}) \tag{1.1}$$

1.1.3 Unit Analysis

Unit analysis is a method used to convert from one unit of measure to another. As previously seen, derived units are obtained from base units in the manner given in Table 1.2 without other numerical factors.

Again, we note that the coulomb is defined as the amount of electric charge carried past a boundary in one second by a current of one ampere. That is, in general, for a constant current I (given in amperes), the amount of charge Q (in coulombs) passing a boundary in a specific amount of time t (in seconds) is determined using equation 1.2.

$$Q = It[\mathrm{C}] \tag{1.2}$$

From this we see that, 2 A will carry 4 C past a boundary in 2 s.

Think About It

How much energy is expended when a force of 85 N pushes an elemental charge over a distance of 5×10^{-3} m?

The unit of energy is the Joule which is a derived unit. Looking at Table 1.2 we see that the joule is derived from a combination of force and distance, i.e.

$$\text{energy}(J) = \text{force}(N) \cdot \text{distance}(m) \tag{1.3}$$

Therefore, the energy expended is:

$$\text{energy} = (85)(5 \times 10^{-3}) = 0.425[\text{J}] \tag{1.4}$$

1.1.4 Scientific versus Engineering Notation

Many electrical computations are performed on vary large and/or very small quantities. To facilitate writing these numbers we use a power of 10 notation. There are two main notation schemes for representing numbers.

Scientific Notation is a form of mathematical notation that allows for only one number to be written to the left of the decimal place and then multiplied by any power of 10 to restore the number to its original value. It is a common shorthand notation for writing large or small mathematical values. For example 4,500 g would be written as 4.5×10^3 g in scientific notation. Likewise, 0.04982 A is written as 4.982×10^{-2} A in scientific notation. Note that although 49.82×10^{-3} represents the same number it is not in scientific notation.

Engineering Notation uses prefixes to represent certain powers of ten (see Table 1.3). There is a significant convenience to using engineering prefixes, however, you cannot forget when doing calculations that engineering prefixes each have a numerical value. To write a value in engineering notation, you write the number using power of 10 notation, where the power is a multiple of three, then use the prefix name for that power of ten. For example, 0.04982 A is written in engineering notation as 49.82 mA.

Table 1.3: Table of Engineering units.

Prefix	**Power of 10**	**Abbreviation**
atto	10^{-18}	a
femto	10^{-15}	f
pico	10^{-12}	p
nano	10^{-9}	n
micro	10^{-6}	μ
milli	10^{-3}	m
kilo	10^{3}	k
mega	10^{6}	M
giga	10^{9}	G
tera	10^{12}	T
peta	10^{15}	P
exa	10^{18}	E

Think About It

A mega-joule (MJ) is equivalent to one 1×10^6[J] $= 1,000,000$ [J].

A micro amp, (μA), is equivalent to 1×10^{-6} [A] $= 0.000001$ [A].

2 mW = 2 milliwatts = $2 \times 10^{-3} = 0.002$ W

5 μC = 5 micro-coulombs = 5×10^{-6} C = 0.000005 C

0.000057 ampere = 57×10^{-6} A = 57 μA

How many different ways can 452,300,001.0 W be written in engineering notation?

- 452,300,001,000,000,000,000,000,000. aW
- 452,300,001,000,000,000,000,000. fW
- 452,300,001,000,000,000,000. pW
- 452,300,001,000,000,000. nW
- 452,300,001,000,000. μW
- 452,300,001,000. mW
- 452,300.001 kW
- 452.300001 MW
- 0.452300001 GW
- 0.000452300001 TW
- 0.000000452300001 PW
- 0.000000000452300001 EW

1.2 Representing Numbers

Some numbers are exact. For instance countable objects like 2 apples. Fractions can also be exact, 1/2 is exactly equal to 0.5. Some numbers are inexact. For example, π has an infinite number of digits. How many digits we choose depends on how close to the exact value we wish to be. Measured values are invariably inexact. A length of wire can only be measured to the accuracy of the ruler used. A resistor value is only as accurate as the manufacturing process used.

1.2.1 Significant Digits

Precision of a value is the number of digits in the value that are treated as significant for computation. In other words the precision of a value is the number of significant digits in that value. Significant digits are those digits that carry actual information.

They include:

- All non-zero digits (1-9) are significant.
- Zeros that have any non-zero digit anywhere to the **left** of them are considered significant zeros.
- All other zeros are **not** considered significant.

Think About It

How many significant digits are in the following numbers?

- 67.0034 inches
- 0.0014800 feet
- 1,000,000 meters

67.0034 has six significant digits (four non-zero digits and two significant zeros).

0.0014800 feet has five significant digits (three non-zero digits and two significant zeros).

1,000,000 is ambiguous. It could be exact, in other words have an infinite number of significant digits and we just didn't want to write them all down. Without a decimal point in the number you really don't know how many significant digits to count. We need some rules for determining whether a given number is exact or not.

- Numbers without decimal points are assumed to be exact, i.e. 10 resistors.
- Numbers with decimal points have the number of significant digits determined as outlined above.

Here is 1,000,000 re-written in two different ways. How many significant digits does each have?

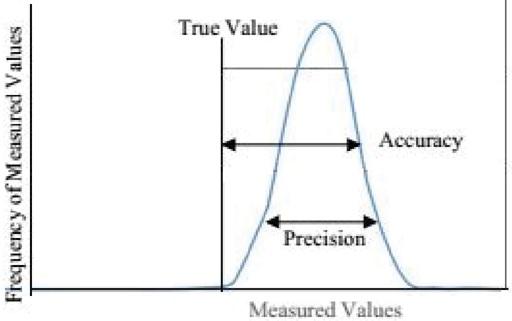
- 1.00×10^6 V
- 1.00000×10^6 V

1.00×10^6 V has three significant digits.
1.00000×10^6 V has six significant digits.

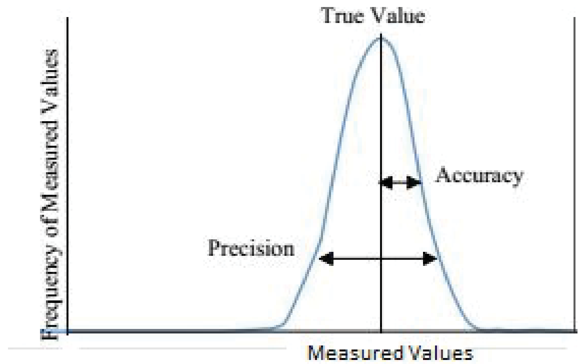
1.2.2 Accuracy and Percision

The accuracy and precision of computed values depends on the accuracy and precision of the numbers used to compute them.

- **Accuracy** is the measure of *absolute error* in a value or measurement.
- **Precision** is a measure of the repeatability in a set of measurements and represents the *relative error* in the value.



(a)



(b)

Figure 1.1: Examples of frequency distribution (Histograms) of measured values showing low accuracy (a) and high accuracy (b). Note that both graphs show the same precision of measurements.

Figure 1.1 shows two histograms for a number of measurements of a particular quantity. In graph (a) the distribution of the measurements are offset from some true value. The offset of the mean of the measurement distribution from the true value is called the measurement bias. Measurement bias can occur due to calibration errors or by not measuring the quantity correctly. In graph (b) the mean of the distribution of the measurements is the same as the true value. Therefore, graph (b) shows no measurement bias.

The accuracy is shown as the difference between a measurement and some true value and includes any measurement bias. Graph (b) shows a better accuracy then graph (a). The precision is how repeatable the measurements are, or the spread of the measurements about some mean. Note that both graphs in Figure 1.1 show measurements with the same distribution or spread about the measurement mean. For this example the two sets of measurements show the same precision. If one of the distributions was narrower than the other, the narrower distribution would have better precision.

You can have precision without accuracy, but you cannot have accuracy without precision! In this class we will assume that our test equipment is calibrated and that the students making the measurements are doing so correctly, so that we won't have any bias measurements. In this case accuracy and precision are somewhat synonymous.

Think About It

You have been asked to assess the accuracy and precision of three different manufacturing processes that are used to make 1.0 kΩ resistors. You measure a sample of resistors from each process (see Table A.)

Table A: Measured values of sample resistors from three different manufacturing processes.

Method 1 (kΩ)	Method 2 (kΩ)	Method 3 (kΩ)
1.0097	2.862987	1.012951
1.0297	2.753858	1.012958
0.9897	2.186057	1.012955
1.0097	2.600967	1.012944

- Which group(s) are the most accurate?
- Which group(s) are the most precise?
- Which group is the most accurate and precise?

Remember accuracy is the measure of absolute error in the measurement. The error is given as the difference between the measurement and the true value, in this case 1.0 kΩ.

Table B: Absolute error determined by $|x_i - m|$ where x_i are the measured values from Table A and m is the true value 1.0 kΩ.

Method 1 (kΩ)	Method 2 (kΩ)	Method 3 (kΩ)
0.0097	1.862987	0.012951
0.0297	1.753858	0.012958
0.0103	1.186057	0.012955
0.0097	1.600967	0.012944

From this table we see that Method 1 and Method 3 have the best accuracy, i.e. the smallest absolute error. Although Method 1 has the best accuracy for individual resistors, it does not have the best accuracy for all resistors. On average Method 1 and Method 3 are similar in terms of accuracy.

One measure of precision might be to take the difference between the largest and smallest measurements in Table A for each Method. Alternative measures for precision might be the variance or standard deviation of the sample.

Table C: Difference between maximum and minimum measurements from Table A.

Method 1 (kΩ)	Method 2 (kΩ)	Method 3 (kΩ)
0.04	0.676930	0.000014

From Table C we see that Method 3 has the best precision although Method 1 precision is not bad.

It is clear that Method 2 has the worse accuracy and precision. Although Method 3 appears to have a bias in it (i.e. all the values are greater than the expected value 1.0 kΩ), it still has the best accuracy and precision of the three methods.

1.2.3 Computation

In order to determine the accuracy and precision of a computed value which was calculated from values with differing number of significant digits we need to identify the value with the fewest significant digits.

- **Multiplication and/or Division-** The product or quotient will be reported as having as many significant digits as the number involved in the operation with the *least* number of significant digits.
- **Addition and/or subtraction-** With all numbers involved in the operation written in the same base unit, the result will have the same number of positions to the right of the decimal as the number with the least number of positions to the right of its decimal.

Think About It

Determine the result to the correct number of significant digits for the following computations:

1. $0.000170 \text{ m} \times 100.40 \text{ m} = ?$
2. $2.000 \times 10^4 \text{ in}^2 \div 6.0 \times 10^{-3} \text{ in} = ?$
3. $(20.04)(16.0)(4.0 \times 10^2) =$
4. $45.621 + 4.3 - 6.61 = ?$
5. $7.342 + 1.54 \times 10^2 = ?$
6. $35.201 + (101 \times 10^{-3} \div 0.2 \times 10^{-2}) - 12.536 = ?$

Solution

1. $0.000170 \text{ m} \times 100.40 \text{ m} = 0.017068 \text{ m}^2 \rightarrow 0.017\text{m}^2$ [Three significant digits.]
2. $2.000 \times 10^4 \text{ in}^2 \div 6.0 \times 10^{-3} \text{ in} = \dfrac{20000.[\text{in}^2]}{0.006[\text{in}]} = 3333333.333 \text{ in} \rightarrow 3.3 \times 10^6 \text{ in}$ [Two significant digits.]
3. $(20.04)(16.0)(4.0 \times 10^2) = 128256 \rightarrow 12. \times 10^4$ [Two significant digits.]
4. $45.621 + 4.3 - 6.61 = 43.511 \rightarrow 43.5$ [Three significant digits.]
5. $7.342 + 1.54 \times 10^2 = 7.342 + 154. = 161.342 \rightarrow 161.$ [Three significant digits.]
6. $35.201 + (101 \times 10^{-3} \div 0.2 \times 10^{-2}) - 12.536 = 35.201 + (\frac{0.101}{0.002}) - 12.536 = 35.201 + (50.5) - 12.536 = 73.165 \rightarrow 73.1$ [Three significant digits.]

1.2.4 Rules for Rounding

When reporting numbers to the correct number of significant digits you will invariably need to round off digits in the calculation. The first step is to determine the last digit that can be considered significant. Consider this numeric value:

$$56.2365496$$

- If the digit to the right of the last reported digit is *less* than 5 round it and all digits to its right off.

$$56.\mathbf{2}\underline{3}65496 \rightarrow 56.2$$

- If the digit to the right of the last reported digit is greater than 5 round it and all digits to its right off and increase the last reported digit by one.

$$56.23\mathbf{6}\underline{5}496 \rightarrow 56.237$$

- If the digit to the right of the last reported digit is a 5 followed by either no other digits or all zeros, round it and all digits to its right off and it the last reported digit is off round up to the next even digit. If the last reported digit is even then leave it be.

$$56.236\mathbf{5}\underline{00} \rightarrow 56.2366$$

  Note: The rules for rounding that you learned in elementary school, i.e. round down if the number to be rounded is less than 5 and round up if the number to be rounded is 5 or greater, unfortunately leads to accumulation of numerical error when dealing with large amount of numbers. In order to handle this, IEEE created the rule that you should round to the nearest even number.

1.2.5 Conversion Factor

Inevitably one has to convert from one type of units to another. Most common are converting from metric units to English units and vis-a-versa. The CRC standard mathematical tables, included at the end of this section, provide conversion factors to 10 significant digits. To obtain the English unit from the metric unit you multiply the metric unit by the conversion factor in the last column. You then report the number of significant digits in the result as equal the number of significant digits in the value with the least number of significant digits.

Using Table 1.4, determine the number of feet to the correct number of significant digits:

1. How many feet in 2.35 meters?
2. How many feet in 2.3 meters?
3. How many feet in 2 meters?

In order to convert meters to feet, multiply the numbers of meters by the conversion factor 3.280839895 (feet/meters). Then determine the number of significant digits using the rules in section 1.2.3.

1. $2.35 \times 3.280839895 = 7.709973753 \rightarrow 3$ significant digits $\rightarrow 7.71$
2. $2.3 \times 3.280839895 = 7.545931758 \rightarrow 2$ significant digits $\rightarrow 7.5$
3. $2 \times 3.280839895 = 6.561679790 \rightarrow 1010$ significant digits $\rightarrow 6.561679790$

Remember that if the numeric value does not have a decimal point, it could be considered exact. Therefore in the last example the number with the least number of significant digits is the conversion factor!

When using conversion factors, you have to be careful in identifying whether the conversion factor itself is exact or not. In Table 1.5 the conversion factors in **bold font** are exact. In other words there are exactly 2.54 centimeters in every inch even without writing an infinite number of zeros after the 4 in the table. When you are working with conversion factors that are exact, you must assume that they have on infinite number of significant digits even if they are not written down.

Table 1.4: **Conversion Factors: Metric - English**
** CRC Standard Mathematical Tables, 27^{th} Ed.

To Obtain	**Multiply**	**By**
inches	centimeters	0.3937007874
feet	meters	3.280839895
yards	meters	1.093613298
miles	kilometers	0.6213711922
ounces	grams	$3.527396195 \times 10^{-2}$
pounds	Kilograms	2.204622622
gallons (U.S. liquid)	liters	0.2641720524
fluid ounces	milliliters (cc)	$3.381402270 \times 10^{-2}$

Table 1.5: **Conversion Factors: English - Metric**
** CRC Standard Mathematical Tables, 27^{th} Ed. ** **Bold** faced numbers are exact

To Obtain	**Multiply**	**By**
centimeters	inches	**2.54**
meters	feet	**0.3048**
meters	yards	**0.9144**
kilometers	miles	**1.609344**
grams	ounces	28.34952313
kilograms	pounds	**0.45359237**
liters	gallons (U.S. liquid)	**3.785411784**
milliliters (cc)	fluid ounces	29.57352956

Using Table 1.5, determine the number of centimeters to the correct number of significant digits:

(1) How many centimeters in 2.354 inches?

(2) How many centimeters in 2.3 inches?

(3) How many centimeters in 2 inches?

In order to convert inches to centimeters, multiply the number of inches by the conversion factor 2.54 (centimeters/inch). Then determine the number of significant digits using the rules in Section 1.2.3.

1. $2.354 \times 2.54 = 5.979160 \rightarrow 3$ significant digits $\rightarrow 5.98$

2. $2.3 \times 2.54 = 5.842000 \rightarrow 2$ significant digits $\rightarrow 5.8$

3. $2 \times 2.54 = 5.979160 \rightarrow \text{EXACT!} \rightarrow 5.08\overline{0}$

In the last calculation both the number of inches and the conversion factor are exact, therefore the result 5.08 is also exact! If we wrote just 5.08 how would we know if it was an exact value or a value with three significant digits? We can write the value as $5.08\overline{0}$ cm, where the over-bar on the zero indicates that the zero repeats forever. Most of the time we will not be dealing with exact values, so this situation does not come up often.

1.3 Waveforms

Frequently electrical parameters, like voltages and currents, are not constant but are time varying. These voltage and current waveforms can take on many different forms. Figure 1.2 shows examples of voltage and current waveforms that we might see in different electrical systems. Obviously sinusoidal waveforms are prevalent. For instance our electric grid uses 60 Hz sinusoidal voltage waveform. Radios and other telecommunications systems also transmit a variety of different sinusoidal frequencies. Square and pulse waveforms are used in control signals, computer signals, clock signals, and the internet for example. The sawtooth waveform is used to control the raster scan or refresh of pixels on your television of computer screen. Triangle waveforms are also used as control signals in various systems.

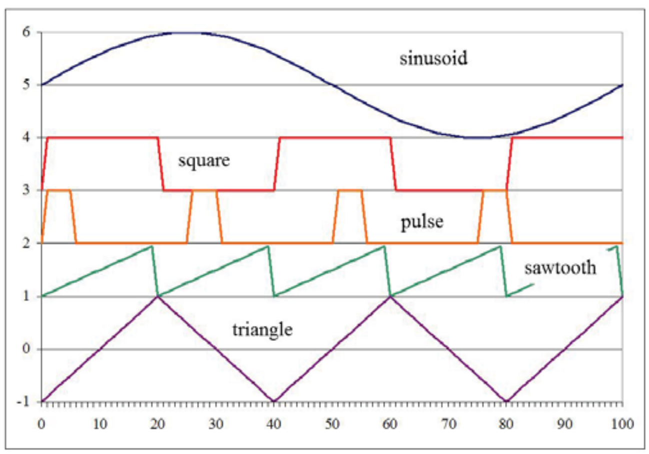


Figure 1.2: Examples of different voltage and current waveforms used in electrical engineering.

In this class we will focus on sinusoidal and pulse waveforms. Later in the curriculum you will learn that any waveform may be decomposed into an infinite series of sinusoids. Therefore any waveform can be described in terms of its sinusoidal decomposition.

1.3.1 Sine Waves

Sine waves are produced by rotating electrical machines such as dynamos and power station turbines; electrical energy is transmitted to the consumer in this form. Figure 1.3 shows an example of a generic voltage sinusoidal waveform. Equation 1.5 describes this waveform mathematically.

$$V(t) = A + B\sin(\omega t + \phi) \tag{1.5}$$

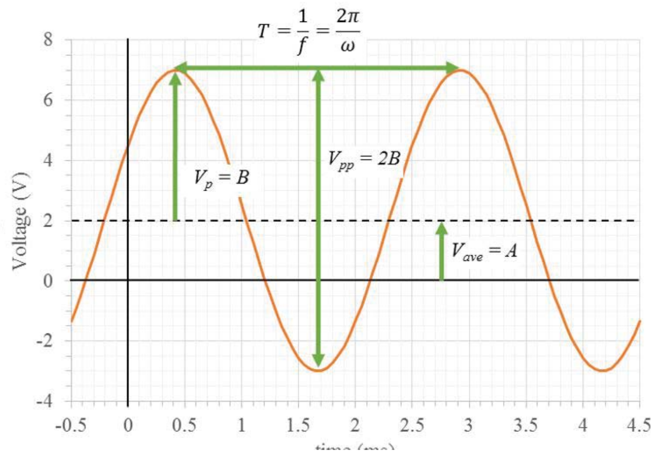


Figure 1.3: Generic sinusoidal waveform.

In order to fully describe a sine wave we need to understand the following terms:

- **Amplitude (B):** The amplitude of the sinusoid is a measure of its change over a single period. It can be specified either as a peak-to-peak amplitude (V_{pp}) or a peak amplitude (V_p) and is measured relative to the average voltage.
- **DC Offset (A):** The DC offset is equal to the time average of the waveform and represents the shift of the waveform up or down on the graph.
- **Period (T):** The period is the time (in seconds) it takes to complete one cycle of a repeating waveform.
- **Frequency (f or ω):** Frequency, f, is defined as the number of cycles per unit time and is the inverse of the period. A related term is radian (or angular) frequency, ω, which is the number of radians per unit time. Since ther are 2π radians per cycle: $\omega = 2\pi f$.
- **Phase (ϕ):** The constant phase (radians) is a measure of the left or right shift of the waveform on the graph.

Each of the parameters in Equation 1.5 : A, B, ω, and ϕ may be estimated directly from the graph of the sinusoid. For example we can determine those parameters for Figure 1.3 as follows.

The amplitude, B, may be determined from the difference of the maximum and minimum voltages:

$$2B = V_{max} - V_{min} \Rightarrow B = \frac{7-(-3)}{2} = 5 \text{ V} \tag{1.6}$$

The DC offset, or average, A, may be determined by taking the average of the maximum and minimum voltages:

$$A = \frac{V_{max} + V_{min}}{2} \Rightarrow A = \frac{7+(-3)}{2} = 2 \text{ V} \tag{1.7}$$

In order to determine the radian frequency, ω, we first need to determine the period, T. The period is determined by the time between two corresponding points on adjacent cycles of the waveform. for the waveform in Figure 1.5, the first maximum occurs at $t_1 = 0.4$ ms and the second at $t_2 = 2.9$ ms. Therefore:

$$T = t_2 - t_1 = 2.9 - 0.4 = 2.5 \text{ ms} \tag{1.8}$$

The radian frequency is therefore:

$$\omega = \frac{2\pi}{0.0025} = 2\pi \cdot 400\text{rad/s} \tag{1.9}$$

We see from this that the frequency, $f = 400$ Hz. Note that a large frequency implies a short period. A small frequency implies a long period.

The last parameter that we need to determine is the phase (ϕ). The phase can be solved for from Equation 1.5 at a specific time. For example, at time $t = 0, \text{v}(0) = 4.5$. Therefore:

$$v(0) = 4.5 = 2 + 5\sin(\phi) \Rightarrow \phi = \sin^{-1}(\frac{4.5-2}{5}) = \frac{\pi}{6} \text{ rads} \tag{1.10}$$

1.3.2 Piece-wise Constant Waveforms

Figure 1.4 shows three different types of piece-wise constant waveforms. Ideal piece-wise constant waveforms have a finite number of levels of various durations and will transition between those levels instantaneously. Square waves (a) have similar features to sine waves in that their maximum and minimum have the same amplitude and duration. Pulse or rectangular waves (b) do not have a symmetrical shape and may not have the same maximum and minimum amplitude. Both square and pulse waves are used in digital circuits (i.e. clock signals, counters, address lines, ect.) Some signaling systems will use multiple levels and can be described as in (c).

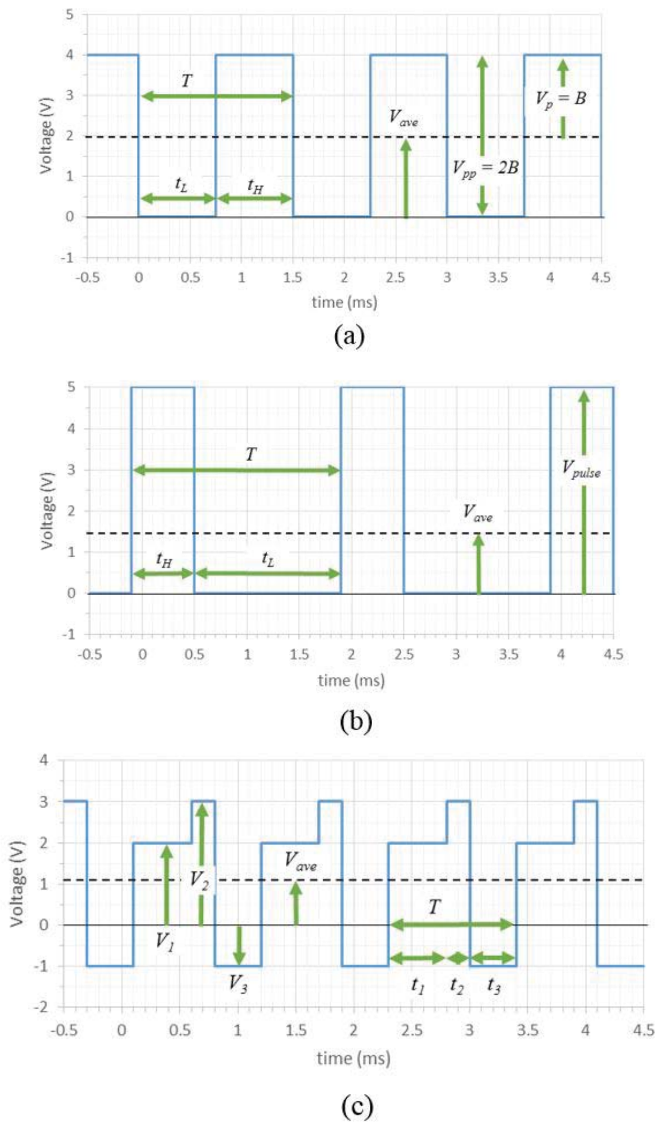


Figure 1.4: Piece-wise constant waveforms: (a) square wave, (b) pulse of rectangular wave, (c) generic piece-wise constant wave.

Some of the terms that describe piece-wise constant waveforms are similar to those that describe sinusoidal waveforms.

- **Amplitude:** Peak amplitude, (V_p), is only unambiguous for a symmetric periodic waves like sine and square waves. V_p is measured relative to V_{ave}. For asymmetric waves one measures the instantaneous amplitude form zero reference. For pulse waveforms this is called V_{pulse}. Strictly speaking this is no longer "amplitude" since there is potentially a DC offset to the waveform. For other generic piece-wise constant waveforms there may be more than one instantaneous value that describes the waveform.

- **Period (T):** The period is the time (in seconds) it takes to complete one cycle of a repeating waveform. For a pulse waveform, this is sometimes called the pulse repetition rate (PRR).

- **Pulse Repetition Frequency (PRF):** The pulse repetition frequency is defined as the number of cycles per unit time and is the inverse of the pulse repetition rate.

- **Duty Cycle (D):** The duty cycle of a pulse or rectangular wave is the percentage of time the pulse is high compared to the total period. A square wave can be considered a pulse wave with a 50% duty cycle.

Square wave:
Following the same process as for a sinusoidal wave, the square wave amplitude and average is determined by the maximum and minimum values. For Figure 1.4 (a) both the amplitude and average voltage are:

$$2B = V_{max} - V_{min} \Rightarrow B = \frac{4-(0)}{2} = 2 \text{ V} \tag{1.11}$$

$$V_{ave} = \frac{V_{max} + V_{min}}{2} \Rightarrow V_{ave} = \frac{4+0}{2} = 2 \text{ V} \tag{1.12}$$

The period or pulse repetition rate is read directly off the graph. For Figure 1.4 (a) the period is 1.5 ms. The pulse repetition frequency is the inverse of the pulse repetition rate and is PRF=666.67 Hz.

Pulse wave:
This is a asymmetric waveform, so we measure the instantaneous amplitude relative to a zero reference. Pulse waveforms only have two levels and one of them is always zero, so frequently we only specify the non-zero level. For Figure 1.4(b) this level is $V_{pulse} = 5$ V. Pulse repetition rate is read directly off the graph, PRR=2.0 ms. Again the PRF is the inverse of the PRR so PRF=500 Hz. The shape of the pulse waveform is described by the duty cycle. Duty cycle is the time high relative to the total period and is typically given as a percentage.

$$D = \frac{t_H}{T} \cdot 100\% \Rightarrow D = \frac{0.6}{2} \cdot 100\% = 30\% \tag{1.13}$$

There are several ways to think about how to determine the average of a waveform. One way is to think about the waveform as a stack of tricks that you need to level off. You really only need to determine the average over one period since all other periods or the waveform are identical. In Figure 1.5 (a) we show two periods of the pulse waveform from Figure 1.4 (b). All the "bricks" are stacked in two towers six bricks wide and 10 bricks tall. We remove bricks form the top of the

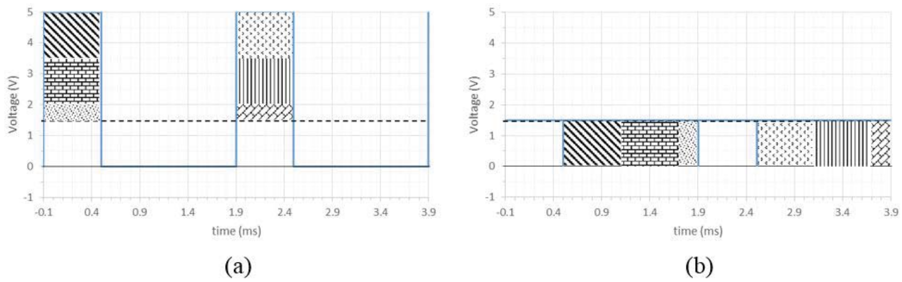


Figure 1.5: Brick by brick interpretation of average.

stacks to fill in in-between the stacks until the entire "wall" is level, see Figure 1.5 (b).

An alternative method, more general method, is to determine the area under the curve and then determine the height that distributes that area evenly across the entire period of the waveform. The area under the curve is determined by calculating the area of each rectangle represented by each voltage level. For our original pulse waveform there are two levels, see Figure 1.6 (a), 5 V for a duration of 0.6 ms and 0 V for a duration of 1.4 ms. Remember that the total period is 2.0 ms. So the average voltage is given by:

$$V_{ave} = \frac{V_1 \cdot t_1 + V_2 \cdot t_2}{T} = \frac{(5\ V)(0.6\ ms) + (0\ V)(1.6\ ms)}{2\ ms} = 1.5\ V \tag{1.14}$$

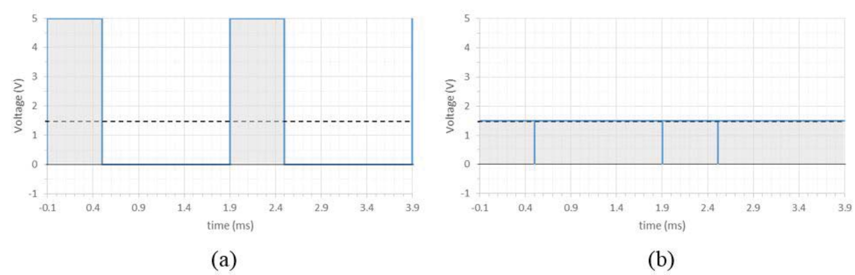


Figure 1.6: Area under the curve.

Piece-wise constant waveform:
As before, the period of any repeating waveform is the time it takes to complete one cycle. Examining Figure 1.4 (c) we determine that the piece-wise continuous waveform has a period of 1.1 ms. For this waveform that are three distinct instantaneous voltage amplitudes measured relative to the zero reference, $V_1 = 2$ V for a duration of $t_1 = 0.5$ ms, $V_2 = 3$ V for a duration of $t_2 = 0.2$ ms, and $V_3 = -1$ V for a duration of t_3 =0.4 ms (see Figure 1.7).

The average voltage is then determined by spreading sum of the areas for each level across the entire period, i.e.:

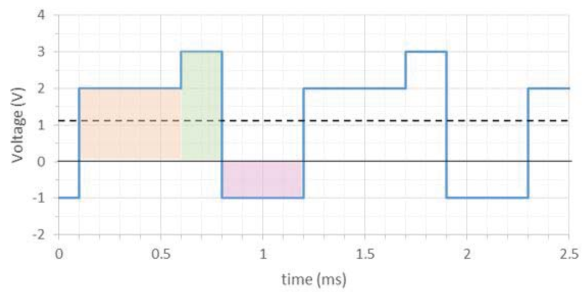


Figure 1.7: Brick by brick interpretation of average.

$$V_{ave} = \frac{V_1 \cdot t_1 + V_2 \cdot t_2 + V_3 \cdot t_3}{T} = \frac{(2\text{ V})(0.5\text{ ms}) + (3\text{ V})(0.2\text{ ms}) + (-1\text{ V})(0.4\text{ ms})}{1.1\text{ ms} = 1.1\text{ V}} \tag{1.15}$$

Chapter 2

Basic Concepts

2.1 Key Terms

- **Coulomb (C)**: One Coulomb is equal to the charge carried by approximately $6.24150948 \times 10^{18}$ elementary charges(electrons or protons).
- **Current (A):** A measure of the amount of positive electric charge passing a point in an electric circuit per unit of time.
- **Voltage (V)**: Is the difference in electric potential energy between two points.
- **Power (W):** Rate of energy dissipation.
- **Energy (J)**: Is the capacity of a physical system to preform work. Electric energy is made available by the flow of electric charge through a conductor.
- **Efficiency (η)**: Is the measure of useful power coming out of a system relative to the total power going into the system. Efficiency is unitless.

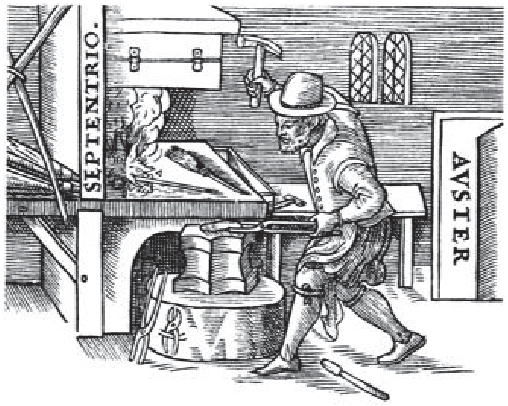
2.2 History

The most primitive electrical and magnetic phenomena, attraction of chaff to rubbed amber and iron to loadstone, has been observed before recorded history. The first application of magnetism, a compass, was used as long ago as 2637 BC by the Emperor Huan-ti of China. The first scientific-type explanation was by Lucretius (98 BC - 55 BC) describing a loadstone attracting a ring of iron:

> It must needs be that there stream off this stone very many seeds or an effluence, which, with its blows, parts asunder all the air which has its place between the stone and iron. When this space is emptied . . . Atoms of the iron start forward and fall into the void, all joined together . . . the ring follows . . . with its whole body.

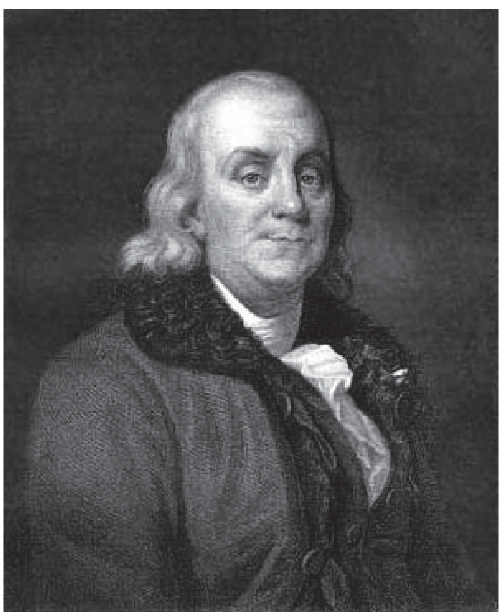
At least this explanation does not invoke gods or other spirits.

William Gilbert (1540-1603)



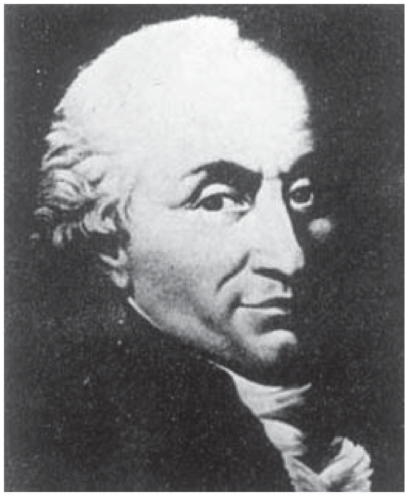
William Gilbert is considered the founding father of the modern science of electricity and magnetism. He was the court physician to Queen Elizabeth of England and published 6 books on electricity and magnetism. He showed clearly how science might be fruitfully pursued through experimentation. Gilbert wrote: “Many modern authors have written about amber and jet attracting chaff and other facts unknown to the generality: with the results of their labors bookseller’s shops are crammed full. Our generation has produced many volumes about recondite, abstruse and occult causes and wonders, and in all of them amber and jet are represented as attracting chaff; but never a proof from experiment, never a demonstration do you find in them.”

Benjamin Franklin (1706-1790)



Benjamin Franklin was the first American to make major contributions to science. He was the first to theorize that there were two types of electrical charges, positive and negative. During Franklin’s time no one had ever “seen” a charge before. Franklin theorized this based on observations and effect. Franklin also showed that lightning is an electric phenomena. Franklin’s “kite experiment” was written up by Joseph Priestly in 1767 “History and Present Status of Electricity.” Evidence shows that Franklin, if he actually did the experiment, would have been insulated (not in a conducting path) otherwise he would have been electrocuted. Unfortunately, others, such as Prof. Georg Richmann of St. Petersburg, Russia, were spectacularly electrocuted when they tried to duplicate the experiment.

Charles A. Coulomb (1736-1806)



Charles-Augustine de Coulomb began the quantitative study of electricity. He established Coulomb's Law which states that the force of attraction or repulsion between two charged spheres is inversely proportional to the square of the distance between their centers. Coulomb's Law is written:

$$F = k\frac{Q_1 Q_2}{R^2}[\mathrm{N}]$$

where Q_1 and Q_2 are the respective charges, in coulombs, on the two spheres, R is the distance, in meters, between them and $k = 8.987551787 \times 10^9$ is Coulomb's constant. Coulomb's Law shows how like charges ($--$ or $++$) repel each other while unlike charges ($+-$ or $-+$) attract each other.

2.3 Atomic Structure and Materials Classification

The atomic structure of matter affects how easily charges move through a material and hence how the material is used electrically. Atoms (see Figure 2.1) are composed of a central nucleus, which contains protons and neutrons. Orbiting around an atom's nucleus are electrons. Each element has its own unique combination of electrons (e^-), protons (p^+), and neutrons (n^o). For example, copper has 29 electrons, 29 protons, and 35 neutrons. Silicon has 14 electrons, 14 protons, and 14 neutrons.

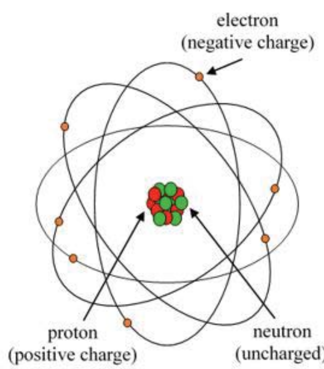


Figure 2.1: Bohr model of an atom.

Table 2.1: Characteristics of electric particles.

Particle	**Charge**	**Mass**
neutron	0	1.67×10^{-27} kg
proton	1.602177×10^{-19}	1.67×10^{-27} kg
electron	$-1.602177 \times 10^{-19}$	9.1×10^{-31} kg

The Bohr model of the atom has electrons moving in spherical orbits called shells. Only a certain number of electrons can exist within a given shell: 2=K, 8=L, 18=M, 32=N. The number of electrons in any shell depends on the element. Copper, for instance, fills the first three shells (2+8+18=28) and has one left over for the N shell. The outer shell is called the valence shell and contains the valence electrons. No element can have more than 8 valence electrons. The number of valence electrons that an element has directly affects its electrical property.

Electrons in the outer orbits are less strongly attracted (less tightly bound) to the nucleus than those in the inner orbits (remember Coulomb's Law). Valence electrons are the least tightly bound and will, if given enough energy, cross the gap between the valence band and conduction band and escape from their parent atom. The amount of energy required depends on the number of electrons in the valence shell. If an atom has only a few valence electrons, only a small amount of energy is needed. For example, copper's valence electrons may have enough energy by heat alone (thermal energy) even at room temperature to escape the parent atom. In fact, at room temperature 1 cm^3 piece of copper has approximately 8.4×10^{22} free electrons, i.e. electrons that ae not bound to the parent atom.

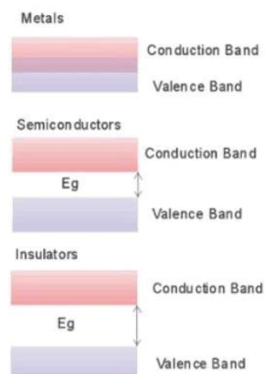


Figure 2.2: Relative position of materials valence band and conduction band with respect to their band gap energy, Eg.

Materials are classified by the amount of energy needed to move their valence electrons from the valence band to the conduction band, where they are free from their parent atom. For metals, such as copper, there are few valence electrons and the valence band and conduction band overlap, hence the large number of charges that are free to move.

Materials that are considered insulators, e.g. glass, porcelain, plastic, have valence shells that are mostly full. It takes a lot of energy Eg, to move valence electrons into the conduction band. It is very difficult to move charges through an insulator. However, if large enough energy is applied the force is so great that the electrons are torn form their parent atom causing insulation breakdown. (An example of insulation breakdown can be found at `http:youtu.be/3M9hfZ47pq8`)

Semiconductors are electrically somewhere between conductors and insulators. Their valence shells are perhaps half full. For example, silicon, fills shells K and L leaving four electrons for its valence shell making it exactly half full. Semiconductors have unique electrical properties which will be discussed further in Chapter 8.

2.4 Charge and Current

The unit of charge is the coulomb (C) and is defined as the charge carried by $6.24150948 \times 10^{18}$ elementary charges (protons or electrons). Inversely, the charge on one elementary particle (proton or electron) is $\pm 1.602177 \times 10^{-19}$.

Think About It:

An initially neutral body has 1.70 µC of negative charge removed. Later, 18.7×10^{11} electrons are added. What is the body's final charge?

Solution:

Initially the body is neutral, thus:

$$Q_i = 0 \text{ C}$$

Next: -1.70 µC is removed:

$$Q_i - (-1.70 \text{ µC}) = +1.70 \text{ µC}$$

is now on the body.

Finally, 18.7×10^{11} electrons are added. Each electron has a charge of $-1.602177 \times 10^{-19}$ C, so

$$(18.7 \times 10^{11})(-1.602177 \times 10^{-19}) = -0.29961 \text{ µC}$$

are added.

Therefore, the final charge is:

$$+1.70 \text{ µC - } 0.29961 \text{ µC} = 1.40 \text{ µC(to 3 significant digits).}$$

Electric Current is the rate of flow of electric charge. Current is defined as the amount of charge that passes a reference point in one second. Equation 2.1 shows that current, I (in amperes), is determined by taking the time derivative of a potentially time, t (in seconds), varying charge, Q (in Coulombs), passing a reference point. If the rate of flow is uniform then the time derivative reduces to a calculation of the amount of charge passing a reference in a specified amount of time. In this class we will always assume a constant or uniform rate of flow of charge.

$$I = \frac{dQ}{dt} = \frac{\Delta Q}{\Delta t} \text{(uniform flow)} \tag{2.1}$$

Strictly speaking it is incorrect to say that current 'flows' or 'goes' form one point to another. It is the charge that flows. Current is the measure of that flow. Additionally, note that current is defined as the rate of flow of *positive* electric charge. At the time that current was defined it was not known that it is the electrons that actually move. The net result is that the direction of current is opposite the flow of electrons.

Remember that in a conductor, like a bar of copper, there are lots of electrons that are not bound to their parent atom and are free to move (see Figure 2.3(a)). Without a "source" these electrons move randomly in all directions. In other words the bulk motion of electrons is zero and hence the current is zero.

When a source is attached (Figure 2.3(b)), the electrons are attracted to the + terminal of the source and repulsed from the − terminal of the source. The source drives the electrons in a preferential direction.

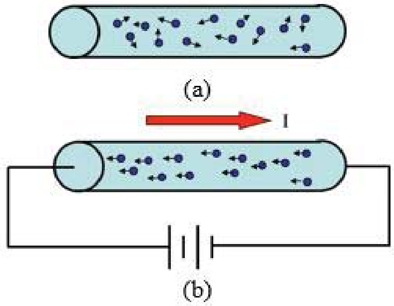


Figure 2.3: Example of a conductor with (a) random movement of electrons and (b) uniform movement of electrons.

Remember that the direction of the current is opposite the direction of the motion of electrons.

To fully specify the current, you must include both its value and its direction. Figure 2.4 shows two identical circuits. The electrons move in the same direction, from the negative terminal of the source around the circuit towards the positive terminal of the source. You can think of the current as the measurement of the positive charge moving from one point to another is a specific direction. So in Figure 2.4(a) the measured current direction is the same as the motion of positive charge (i.e. opposite movement of negative charge) and therefore is positive. In Figure 2.4(b) the measured current direction is opposite the motion of the positive charge (i.e. in the same direction as the negative charge) and therefore is negative. Pay attention to the fact that the circuits are identical. It is only the measurement of the current that changed.

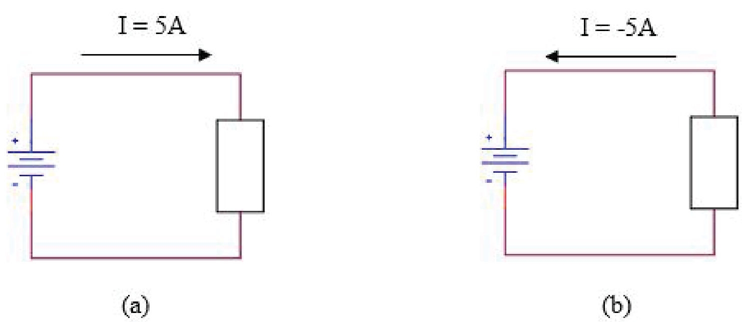


Figure 2.4: Identical circuits with current measured in opposite directions.

Think About It:
(1.) The current in a certain conductor is 40 mA.

a. Find the total charge that passes through the conductor in 1.5 s.

b. Find the total number of electrons that pass through the conductor in that time.

Solution:

a. Total Charge:

$$Q = I \times t = (40 \times 10^{-3}\text{ A}) \times (1.5\text{ s}) = 0.06\text{ C}$$

b. Total number of electrons:

$$\#e = (0.06 \text{ C})(6.24150948 \times 10^{18} \#\text{e/C}) = 3.7 \times 10^{17} \text{electrons}$$

(2.) If 840 coulombs of charge pass a point during a time of 2 minuets, what is the current at that point?

Solution:

$$\frac{\text{charge (C)}}{\text{time (s)}} = \frac{840 \text{ C}}{120 \text{ s}} = 7 \text{ A}$$

Dividing our charge of 840 C by 120 s gives us the rate that the electrons are flowing, thus the current at that point is equal to 7 A.

This is an example of **uniform current**; total amount of charge in coulombs over time in seconds.

2.5 Voltage

Voltage is the difference in potential energy between two points, A and B, and describes the amount of energy required to move a unit charge from point A to point B. Voltage between two points is one volt if it requires one joule of energy to move one coulomb of charge form one point to another.

$$V_{AB} = \frac{E_{AB}}{Q} \Rightarrow \text{volt} = \frac{\text{joule}}{\text{coulomb}} \tag{2.2}$$

In Equation 2.2 the subscripts explicitly indicate that the voltage is measured at node A with respect to node B. In other words the + terminal of the multimeter is connected to node A and the − terminal of the multimeter is connected to node B (see Figure 2.5).

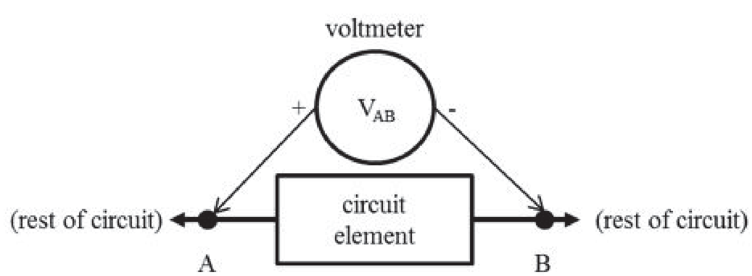


Figure 2.5: Measuring differential voltage

Frequently voltage in a circuit is measured relative to a common node or ground. By definition we assume the potential at the common node is zero. The common node or zero point is somewhat arbitrary and set for convenience. It is the potential difference (voltage) *between* two nodes that has physical meaning. Figure 2.6 shows two different symbols typically used to mean "ground". The symbol shown in Figure 2.6(a) is used for the circuit common node. The symbol shown in Figure 2.6(b) is used to indicate the actual earth ground.

It is frequently useful to measure all voltages relative to the same common node or ground. These measurements are called node voltages. Node voltages are typically subscripted with only the single non-ground node. For example in Figure 2.7 the measurement V_A is a node voltage measured at node A with respect to the circuit common or ground node. V_B is also a node voltage measured at node B with respect to ground.

Figure 2.6: (a) circuit common, (b) earth ground

The measurement V_{AB} is not a node voltage since it is not measured with respect to the circuit ground. The differential voltage V_{AB} can be determined from the node voltage measurements as follows:

$$V_{AB} = V_A - V_B \tag{2.3}$$

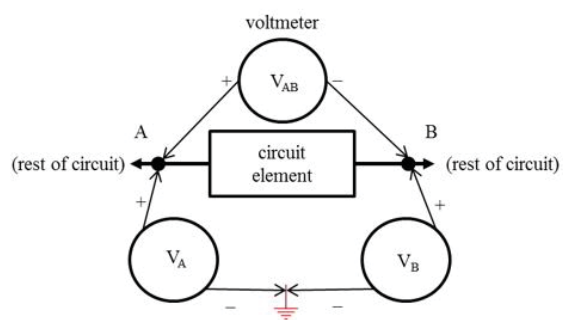


Figure 2.7: Measurement of differential versus node voltages

Things to remember:

- It is not possible to obtain a current in a component connected to nodes A and B unless there exists a potential difference between those two points.

No potential difference $\Rightarrow$ no current
Caution: no current $\nRightarrow$ no potential difference

- Two terminals must always be specified (or understood) when referring to the voltage of a circuit element.
- All points on an ideal wire have the same voltage.
- In a circuit all "grounded" nodes are connected together.

2.5.1 Batteries

Voltage is created by the separation of charge (think static electricity). In a battery (see Figure 2.8) the charges are separated by chemical reaction. An ordinary flashlight battery has an inner carbon rod and an outer zinc case. The chemical reaction between the ammonium-chloride/manganese-dioxide paste and the zinc case creates an excess of electrons; hence the zinc carries a negative charge and is called the negative electrode or anode. An alternate reaction leaves the carbon rod (positive electrode or cathode) with a deficiency of electrons, causing it to be positively charged. This results in an electric potential difference, or voltage, between the anode and the cathode. You can think of this difference as an unstable build-up of electrons. The electrons want to rearrange themselves to get rid of this difference, but they do this in a certain way.

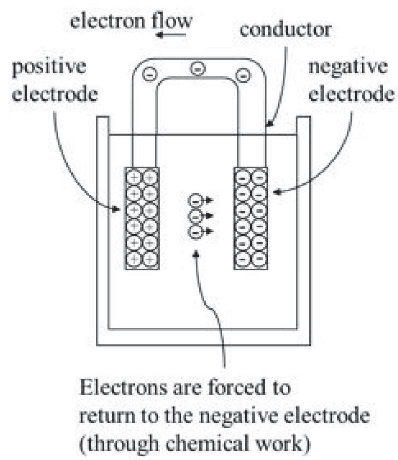


Figure 2.8: Basic diagram of a battery and its electron flow.

In a battery, the only place for the electrons to go is to the cathode. However, the electrolyte keeps the electrons from going straight from the anode to the cathode within the battery. When the circuit is closed (for example a light connecting the anode and cathode) the electrons will have a path to the cathode. In Figure 2.8 the electrons go through the wire. This is one way of describing how electrical potential causes electrons to flow through the circuit. The electrochemical processes change the chemicals in anode and cathode to make them stop supplying electrons, so there is a limited amount of energy available in a battery.

When you *recharge* a battery, you change the direction of the flow of electrons using another power source, such as solar panels, or a second car battery. The electrochemical processes happen in reverse, and the anode and cathode are restored to their original state and can again provide energy.

The chemical reaction that is producing the electrons does not happen instantaneously. The speed of electron production by this chemical reaction controls the rate of electrons that can flow between the terminals, i.e. it limits the maximum current that the battery can supply. We can model this effect by assuming that the battery has some *internal resistance* that limits the maximum current. electrons flow from the battery through a circuit and must travel from the negative terminal back to the positive terminal for the chemical reaction to take place. Therefore a battery can sit unconnected on the shelf for a long period of time and still have plenty of energy when it is used. Once the circuit is connected so that charge is allowed to flow, then the reaction starts.

One of a battery's terminals is always positive and the other is always negative, therefore current flow is always unidirectional in a battery. This unidirectional current is referred to as DC or direct current, and the battery is referred to as a DC source. Alternately, an AC or alternating current is a current that changes direction cyclically; therefore charges will flow in both directions. Your home is powered by an alternating current if it receives its energy from a commercial AC power system.

Batteries are rated by the voltage (V) they provide and the current capacity (mAh) or energy (Wh) available (see Table 2.2 and Table 2.2). A rating of mAh or Wh (capacity) is the number of mA or W that the battery can produce (drain) in one hour (lifetime). The mAh (Wh) rating tells you the theoretical lifetime of the battery at a given load. In other words, lifetime equals capacity divided by drain. A battery has a fixed lifetime. The greater the current, or power delivered, the shorter the lifetime.

$$\text{life[h]} = \frac{\text{capacity[mAh]}}{\text{drain[mA]}} = \frac{\text{capacity[Wh]}}{\text{drain[W]}} \tag{2.4}$$

Table 2.2: Battery Technology

Cell Type	**Nominal Voltage**	**Specific Energy (Wh/kg)**	**Specific Power (W/kg)**
Alkaline[1]	1.5 V	150	14
Lead-acid[2]	2.0 V	35	400
Nickel-Cadmium[2]	1.2 V	80	300

[1] Not rechargeable, [2]Rechargeable

Table 2.3: Size and energy available for standard batteries.

Battery Size	**Energy Available**	
	Alkaline	Ni-Cd
AAA	0.5 Wh	0.2 Wh
AA	1.5 Wh	0.36 Wh
C	5.0 Wh	1.2 Wh
D	15.0 Wh	3.6 Wh

Think About It:
A battery rated at 1400.0 mAh supplies 28.0 mA to a load. How long can it be expected to last?

Solution :

$$\text{life(h)} = \frac{\text{capacity[mAh]}}{\text{drain [mA]}} = \frac{1400.0 \text{ mAh}}{28.0 \text{ mA}} = 50.0 \text{ h}$$

2.6 Power and Energy

Power is defined as the rate of energy *dissipation* or the rate at which work is performed. More accurately the rate of energy conversion, i.e. to heat, light, mechanical energy, etc. Equation 2.5 shows that power, P (in watts), is determined by taking the time derivative of a potentially time, t (in seconds) varying function of energy, E (in joules). If the energy dissipation in a system is uniform then the time derivative reduces to a calculation of the amount of energy dissipated in a specific amount of time. In this class we will always assume a constant or uniform energy dissipation rate.

$$P = \frac{dE}{dt} = \frac{\Delta E}{\Delta t}(\text{uniform dissipaton rate}) \tag{2.5}$$

One watt is therefore equal to one joule per second. The power dissipated in a circuit element can also be determined by the voltage across the element and the current through it. Remember that the units for voltage can be written as joules per coulomb and the unit for current can be written as coulomb per second. So:

$$P(\text{watts}) = \frac{\Delta E}{\Delta t} = \frac{\text{joules}}{\text{second}} = \frac{\text{joules}}{\text{coulomb}} \times \frac{\text{coulomb}}{\text{seconds}} = VI \tag{2.6}$$

Think About It:
A resistor draws 3.0 A from a 12.0 V battery. How much power does the battery deliver to the resistor?
Solution:

$$\text{Power [W]} = (\text{V})(\text{A}) = (12.0 \text{ V})(3.0 \text{ A}) = 36. \text{ W}$$

Energy is defined as the ability to do work. Work is done by moving charges through a circuit. In a battery a chemical reaction forces a separation of charge. The battery then has electric potential energy, i.e. if connected to a circuit the battery would force the movement of charge through the circuit. Remember that energy (in joules) equals charge (in coulombs) times voltage (in volts).

$$E = VQ \tag{2.7}$$

Energy can also be derived from power. Since power is the time rate of change of energy, energy is the integral of a time varying power, or for constant power dissipation, is equal to power times time.

$$E = \int_0^T P(t)dt = PT \text{ (constant power dissipation)} \tag{2.8}$$

Your utility company charges you for how much energy you use. The unit for energy is a joule which is equal to a watt second. This tends to be a very large number. Rather than charge you per joule, most utility companies will charge you per kilowatt-hour (kWh) which is still a measure of energy (like Celsius and Fahrenheit are both measures of temperature).

Think About It:

(1) A resistor draws 24.0 mA from a 120 V battery. How much power does the battery deliver to the resistor?

(2) An electronic oven requires 220 V, 10 A. What is the power requirement?

(3) If the power company charges 10¢ per kWh, how much does it cost to bake a turkey for 5 hs?

Solution:

(1)

$$Power[W] = Voltage[V] \times current[A]$$
$$(12.0\ V) \times (0.024\ A) = 0.288\ W$$

(2) $P = V \times I$

$$220\ V \times 10\ A = 2,200\ W\ or\ 2.2\ kW$$

(3) $E = P \times t\ \therefore\ 2.2\ kW \times 5\ hs = 11\ kWh$

$$Cost = kWh \times 0.1\frac{\$}{kWh} = \$1.10$$

Passive Sign Convention

Power measured in a circuit can be positive or negative depending on whether the circuit element is dissipating energy or supplying energy. By convention we have defined positive power as the rate of energy dissipation so P 0 implies that the energy is being dissipated and the circuit element is therefore considered a *load*. If P<0, then the energy is being supplied and the circuit element is considered a *source*. The question is, if we are to calculate power of a circuit element by multiplying a measured voltage by a measured current, how do we measure the voltage and current? Consider Figure 2.9, showing four different ways to measure the voltage and current of a single circuit element. If we measure $V_{AB} > 0$ then $V_{BA} < 0$. Additionally, if we measure $I_{AB} > 0$, then $I_{BA} < 0$. The powers for the circuit element calculated using the measurements shown in Figure 2.9 are tabulated in Table 2.4 . Is the circuit element a *source* or a *load*? At this point I cannot really tell. Table 2.4 shows that depending on how I measure the circuit element I might get P>0 and I might get P<0. We need a consistent way to measure voltage and current so that when we use those values to calculate power we are able to interpret the meaning of the sign of the power. Passive sign convention dictates that if I measure the current going into the positive terminal of the voltage measurement,

Table 2.4: Power calculated using the measured voltages and currents defined in Figure 2.9.

	$I_{AB} > 0$	$I_{BA} < 0$
$V_{AB} > 0$	P>0	P<0
$V_{BA} < 0$	P>0	P<0

and the power calculated is positive then I have positive energy dissipation and the element is a load. If the calculated power is negative then I have negative energy dissipation or energy generation and the element is a source.

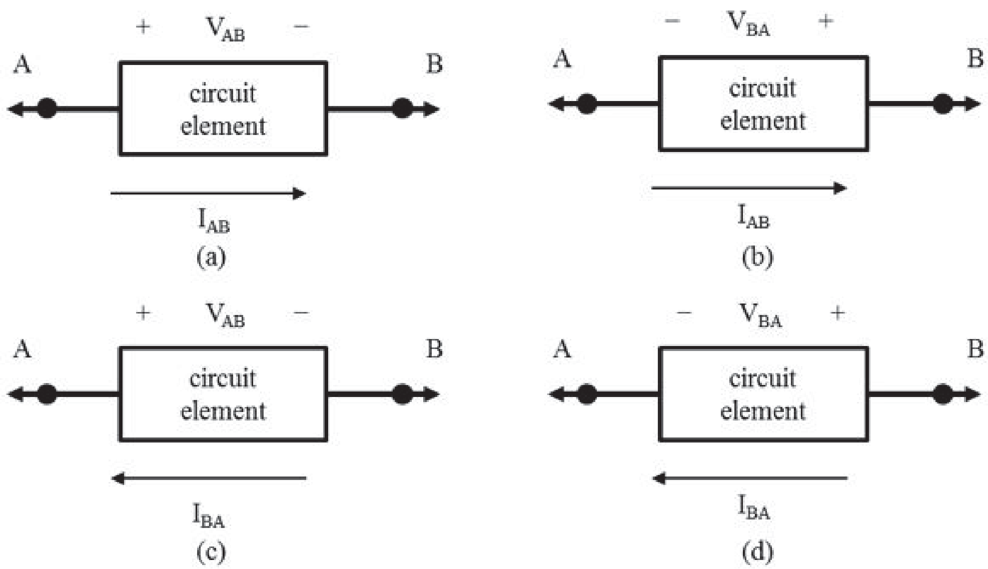


Figure 2.9: Four different ways to measure the voltage and current of a single identical circuit element.

Think About It

Consider the voltage and current measurements shown in Figure 2.9. Which are measured using passive sign convention and which are not?

Solution

Circuit (a) and (d) are measured using passive sign convention.

Consider the circuit measurements shown in Figure 2.10 . The voltage measured across each element is the same since the positive and negative terminals are connected to the same points. The current measurements are shown in passive sign convention with the voltage measurements. The current shown going into element 2 is measured in the opposite direction as the current shown going into element 1. Therefore these two currents are negatives of each other. If we assume that I> 0 and V> 0, which element is a load and which is a source? If we calculate the power for element 2 we find P_2 =VI> 0 is positive and therefore element 2 is a load. For element 1, P_1 =V(-I)< 0 is negative therefore element 1 is a source. Note that the total power in the circuit is zero! For any complete circuit, the power dissipated in the circuit must be equal to the power sourced in the circuit, or equivalently the sum of powers in a circuit must be zero.

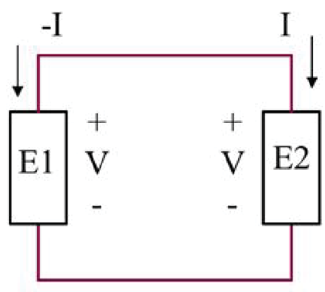


Figure 2.10: Two element circuit.

In complex circuits it is not always apparent which circuit elements are loads and which are sources. Passive sign convention helps make this determination.

2.6.1 Efficiency

Efficiency(η) is a measure of the amount of useful power coming out of a system relative to the total power going into the system. The more "losses" in the system the less efficient the system is. Efficiency is unitless, but can be expressed as either a ratio of energy or a ratio of power.

$$\eta = \frac{P_{out}}{P_{in}} = \frac{P_{out}}{P_{out} + P_{loss}} \tag{2.9}$$

In Figure 2.11, the electrical system has a certain amount of input power. This input power will be converted into usable output power and un-usable "lost" power. The power lost is usually power that is converted into heat and not really destroyed per se, only un-usable. On the other hand if heat is what you are after then the power that was converted into heat would be the usable power. Remember that "lost" power is a matter of definition since the law of conservation of energy (first law of thermodynamics) says that energy (and therefore power) can be converted from one form to another, but it cannot be created or destroyed. Since energy (and power) are conserved the total power of the system is actually zero. In other words $P_{in} = P_{out} + P_{loss}$.

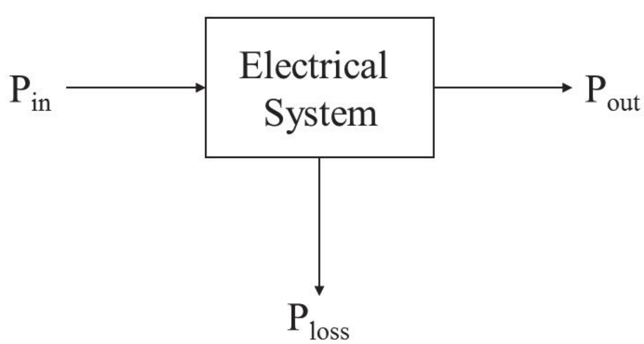


Figure 2.11: Schematic of an electrical system showing power flow.

Think About It:

Let's compare the energy efficiency of incandescent and compact fluorescent light bulbs. The Desired P_{out} is light measured in luminous flux. The P_{loss} is in heat. Table 2.6.1 shows the luminous energy efficiency for each bulb. To be fair, we compare two bulbs with similar luminous output. We first need to convert lumins to watts using the conversion 683 lm = 1 W. Then calculate the luminous efficiency by taking the ratio of P_{out} to P_{in}. Table shows that the compact fluorescent bulb is about four times more efficient than the incandescent bulb.

Table 2.6.1: Comparison of the luminous efficiency for incandescent and compact fluorescent light bulbs.

	Incandescent Bulb	**Compact Fluorescent**
$\mathbf{P_{in}}$ **= Rated Wattage**	60 W	15 W
$\mathbf{P_{out}}$**(683 lm = 1 W)**	$P_{out}[\text{W}] = \frac{X[\text{lm}]}{683\ [\text{lm/W}]}$	
	1.25 W	1.32 W
"Luminous Efficiency"(η)	$\eta = \frac{P_{out}}{P_{in}} \times 100\%$	
	2.1%	8.8%

Efficiency is great, but compact fluorescent bulbs cost more than incandescent bulbs. So which is cheaper to operated over the long run? A typical incandescent bulb has an average lifetime of about 1000 hours. That means that you can operate that bulb about 3 hours/day for an entire year without having to replace it. A typical compact fluorescent bulb has an average lifetime of about 8000 hours and so it will last approximately 8 years at 3 hours/day. Table 2.6.2 shows the cost comparison for operating an incandescent and compact fluorescent bulbs at 1000 hours/year for eight years. In that time you would need to buy eight incandescent bulbs and only one compact fluorescent bulb. So even though the compact fluorescent bulbs cost more, they last longer and are more energy efficient. So over an eight year period you would save $32 for every incandescent bulb you change to a compact fluorescent bulb.

Table 2.6.2: Cost comparison for operating an incandescent and compact fluorescent bulb for 8 years.

	Incandescent Bulb	**Compact Fluorescent Bulb**
Hours used per year	1000	1000
Average lifetime	1000	8000
Cost per bulb*	$1	$9
Price of electricity*	$0.10/kWhr	
Power rating	60 W	15 W
Yearly electricity cost	$6	$1.50
Cost over 10 years	$56	$24

*Price determined in 2014 for Fairbanks, AK

Think About It:

A 120. V dc motor draws 12. A and develops an output power of 1.6 hp. What is its efficiency? How much power is wasted? (Note: there is 0.745712 kW per hp)

Solution:

$$P_{in} = VI = (120)(12) = 1440\ W$$

and

$$P_{out} = 1.6hp \cdot 746\ \frac{W}{hp} = 1194\ W$$

therefore

$$\eta = \frac{P_{out}}{P_{in}} = \frac{hp \cdot \frac{W}{hp}}{VI} = \frac{(1.6\ hp)(745.712\ \frac{W}{hp})}{(120.\ V)(12.\ A)} \times 100 =$$

$$82.9 \Rightarrow 83\%$$

$$P_{loss} = P_{in} - P_{out}$$

$$= 1440. - 1193.139 = 246.8608\ W \Rightarrow 0.25\ kW$$

Chapter 3

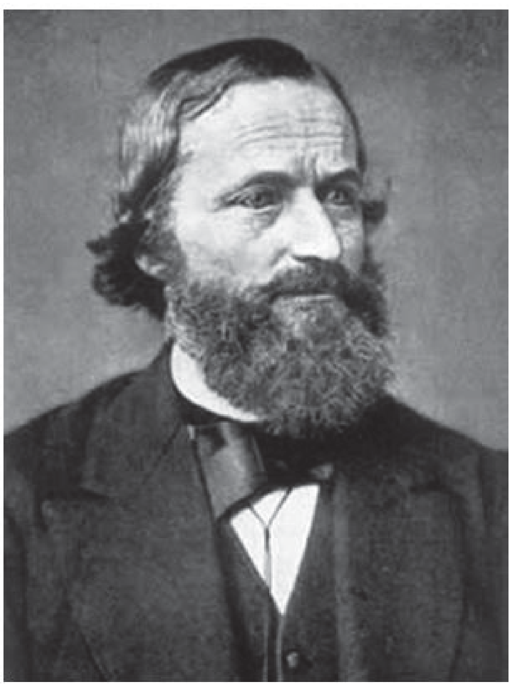
Basic Laws

3.1 Key Terms

- **Kirchhoff's Current Law (KCL):** The algebraic sum of currents entering a node or closed boundary is zero.
- **Kirchhoff's Voltage Law (KVL):** The algebraic sum of voltages around a closed loop is zero.

3.2 History

Gustav Robert Kirchhoff (1824-1887)



Gustav Kirchhoff formulated Kirchhoff's laws in 1845, while he was still a student, extending the theory of the German physicist Georg Simon Ohm. Kirchhoff's laws, along with Ohm's law, form the basis of circuit theory.

Kirchhoff is famous among engineers, chemists, and physicists. In 1854, he worked with Bunsen to found spectrum analysis. Bunsen was analyzing the colors given off by heating chemicals to incandescence using colored glass to distinguish between similar shades. Kirchhoff suggested that the observation of spectral lines, by dispersing the light with a prism, would be a more precise way of testing the color of light. Kirchhoff and Bunsen found that each substance emitted light that had its own unique pattern of spectral lines – a discovery that began the spectroscopic method of chemical analysis. Kirchhoff applied spectrum analysis to the study of the composition of the Sun and was the first to explain that the dark lines in the Sun's spectrum were caused by the absorption of particular wavelengths as the light passes through a gas. Kirchhoff is also credited with the Kirchhoff's law of radiation.

3.3 Branches, Nodes, and Loops

Before we can talk specifically about Kirchhoff's laws we need a few definitions.

- **Branch**: represents a single element such as a voltage source or resistor.
- **Node**: is a point of connection between two or more branches.
- **Loop**: is any closed path in a circuit.

Consider the circuit shown in Figure 3.1. In this circuit there are six (6) **branches**: V1, R1, R2, R3, R4, and I1. There are four (4) **nodes**: the connections between V1 and R1; the connection between R1, R2, R3, and I1; the connection between I1, R3, R4: and the connection between V1, R2, and R4. There are six (6) **loops**. Loop 1 contains V1, R1, and R2. Loop 2 contains R3 and I1. Loop 3 contains R2, I1, and R4. Loop 4 contains V1, R1, I1, and R4. Loop 5 contains R2, R3, and R4. Loop 6 contains V1, R1, R3, and R4. Even though technically there are six loops, they are not all independent, i.e. loops that contain unique information. For most circuits we only need to identify *independent* loops.

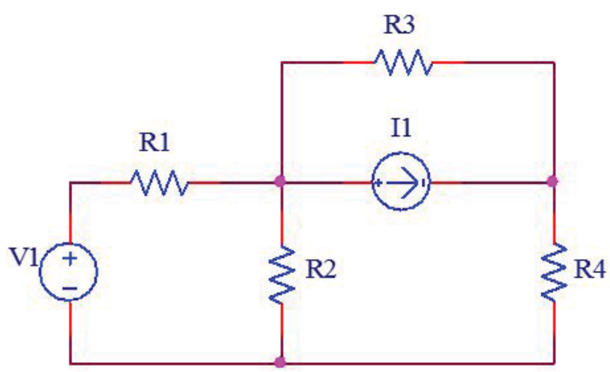


Figure 3.1: Example circuit

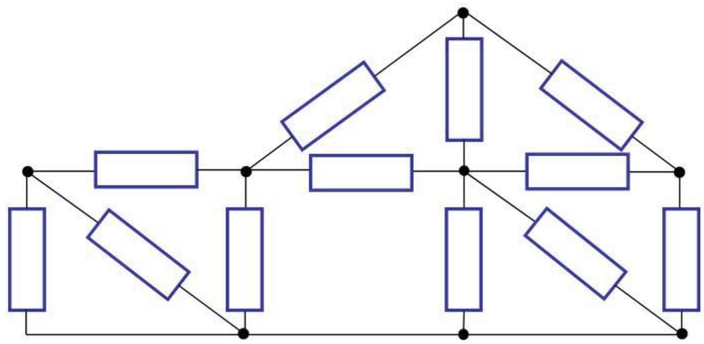
The fundamental theorem of network topology states that a network of b branches, n nodes, and l independent loops will satisfy the following equation:

$$l = b - n + 1 \tag{3.1}$$

So for our example circuit in Figure 3.1 there are $l = 6 - 4 + 1 = 3$ independent loops.

Think About It

How many branches, nodes, and *independent* loops are in the circuit shown?



Solution

Branches = 12, Nodes = 6.
Using Equation 3.1, we can determine the number of independent loops to be (7): $12 - 6 + 1 = 7$.

3.4 Kirchhoff's Current Law (KCL)

Kirchhoff's Current Law (KCL) is based on the conservation of charge which states that the algebraic sum of charges in a system cannot change. Since current is the flow of charge, based on the conservation of charge, the total current going into a node or portion of a circuit must equal the total current coming out of that node or portion of the circuit. In other words there is no accumulation of charge in a conductor. Mathematically we can state Kirchhoff's Current Law as follows:

$$\Sigma\ I_{in} = 0 \tag{3.2}$$

When written this way, we need to pay attention to the sign of the current. Consider Figure 3.2. We would like to write KCL for the node shown. In order to write KCL for this node using equation 3.2, I need all currents measured *entering* the node. I2 and I4 are exiting the node. I need to change the direction of my measurement which changes my measurement to a negative value. Therefore, KCL written at that node is I1-I2+I3-I4+I5=0.

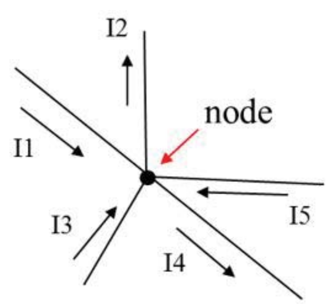


Figure 3.2: Example node illustrating KCL.

An alternative way to write KCL is that the sum of the currents entering a node or closed boundary equals the sum of the currents exiting the node or closed boundary.

$$\Sigma\ I_{in} = \Sigma\ I_{out} \tag{3.3}$$

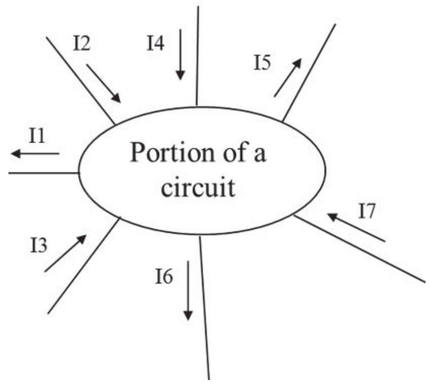


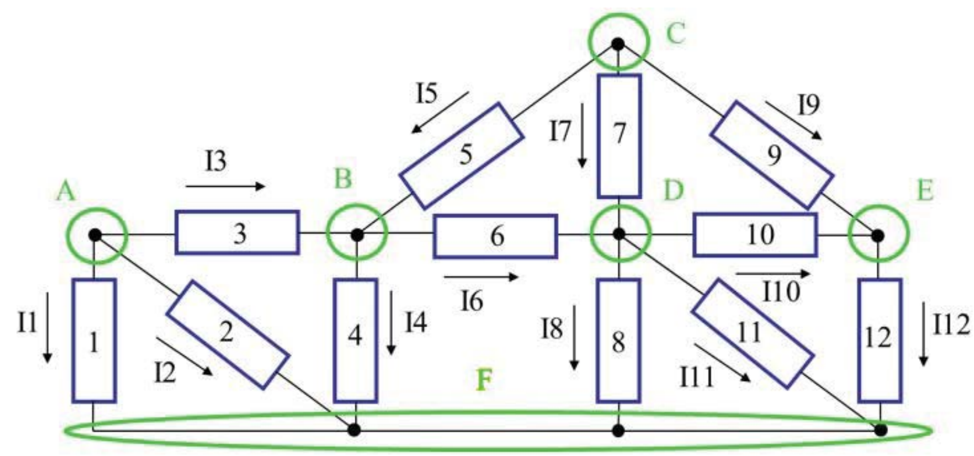
Figure 3.3: Example closed boundary illustrating KCL.

Consider Figure 3.3. I2, I3, I4, and I7 are entering the boundary, while I1, I5, and I6 are exiting the boundary. Therefore, writing KCL for this closed boundary using Equation 3.3 gives, I2+I3+I4+I7 = I1+I5+I6.

Note that both Equations 3.2 and Equation 3.3 are equivalent.

Think About It

Write KCL for each of the nodes A-F shown.



Node A: $I_1+I_2+I_3=0$
Node B: $I_4+I_6=I_5$
Node C: $I_5+I_7+I_9=0$
Node D: $I_6+I_7=I_8+I_{10}+I_{11}$
Node E: $I_9+I_{10}=I_{12}$
Node F: $I_1+I_2+I_4+I_8+I_{11}+I_{12}=0$

3.5 Kirchhoff's Voltage Law (KVL)

Kirchhoff's voltage Law (KVL) is based on the conservation of energy which states that the energy supplied in a circuit must be dissipated in the circuit. Since voltage is a measure of the energy required to move a charge from point A to point B, the conservation of energy implies that if the charge is moved around a closed loop, coming back to its starting point, the energy used and therefore the sum of the voltages equals zero. Mathematically we can state Kirchhoff's Voltage Law as the algebraic sum of voltages around any closed loop equals zero.

$$\sum_{\text{closed loop}} V_i = 0 \tag{3.4}$$

Just like Kirchhoff's Current Law requires that one pays attention to the direction in which the current is measured; Kirchhoff's Voltage Law requires that one pays attention to the polarity in which the voltage is measured.

Consider the circuit shown in Figure 3.4. There are four voltages measured around this loop. Each voltage is measured across a specific branch in the polarity given. We would like to apply KVL (Equation 3.4) to this loop. It does not matter in which direction you go around the loop. Regardless of direction you will derive equivalent equations.

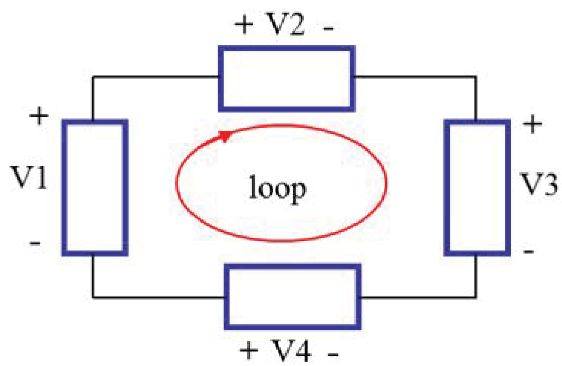


Figure 3.4: Example loop illustrating KVL.

What does matter is how you add up the voltages. One approach is to use as the "sign" of the voltage the measured polarity at the side of the branch that you "enter" as you move around the loop. For example, starting at the lower left-hand corner of Figure 3.4 and moving clockwise around the loop, you enter the negative side of V1, therefore in the KVL equation V1 will be negative. Continuing, we enter the positive side of V2 and V3 and the negative side of V4 before arriving back at the starting point. Therefore the KVL equation around this loop is:

$$-V1 + V2 + V3 - V4 = 0 \tag{3.5}$$

An alternative approach is to consider as you are moving through a branch whether you are traveling from the negative terminal to the positive terminal (a voltage rise) or from the positive terminal to the negative terminal (a voltage drop). If as you move around the loop you pass through a voltage rise, use a positive sign. If you pass through a voltage drop, use a negative sign. So again, examining Figure 3.4, moving clockwise around the loop we travel through a voltage rise in branch 1 (+V1), branch 2 and branch 3 both give voltage drops (−V2, − V3), while branch 4 gives a voltage rise (+V4). Therefore my KVL equation for this loop is:

$$+V1 - V2 - V3 + V4 = 0 \tag{3.6}$$

Equations 3.5 and 3.6 are equivalent. Regardless of which method and direction around the loop you choose, it is essential that you be consistent!

A voltage measurement does not have to be across a particular element or branch. It could be across any arbitrary two points in your circuit. You can still write KVL around the loop that contains that measurement. For example, consider the diagram shown in Figure 3.5. V1 and VA are measured across two arbitrary points in some circuit. V2 is measured across a specific branch or element. VB is measured across an open circuit. Regardless of where the voltages are measured one can still apply KVL around any closed loop. We can define two loops in Figure 3.5, one containing V1 and V2 and the other containing VA and VB. KVL around those two loops gives $-V1 + V2 = 0$ and $VA + AB = 0$.

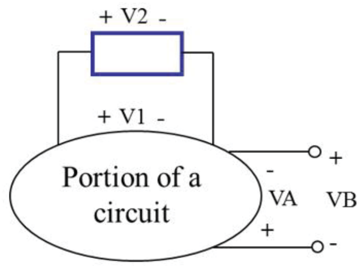
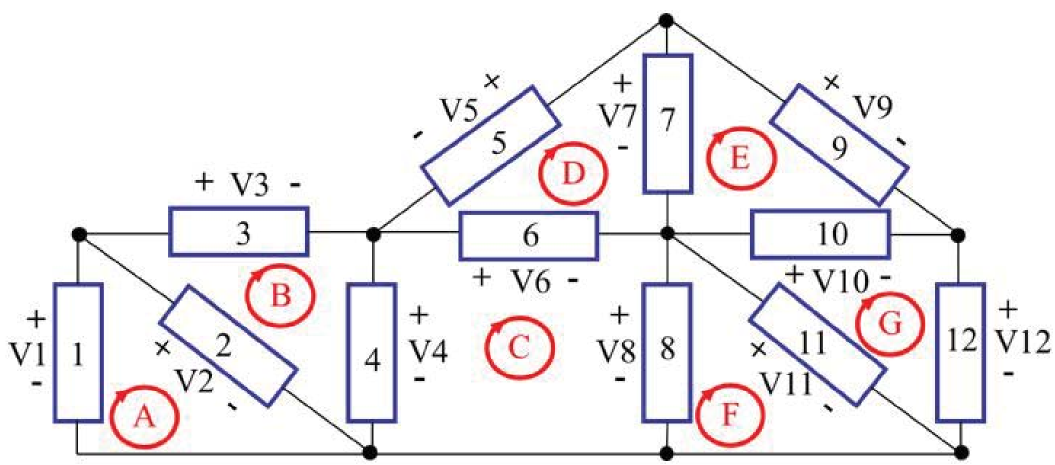


Figure 3.5: KVL around arbitrary voltage measurements.

Notice that in the first loop V1 and V2 were measured in the same direction or polarity. The plus sides of the measurements are on the same point as are the negative sides of the measurements. Your equation shows that these two measurements are equal to each other. These voltage measurements are in “parallel” and measured in the same “polarity”. In the second loop, VA and VB are still in parallel, but are measured in opposite polarity, i.e. the plus side of VA is connected to the same point as the negative side of VB. The KVL equation we derived shows that these two measurements are negatives of each other.

Think About It

Write KVL for each of the loops A-G shown.



Solution

Loop A: $-V_1+V_2=0$

Loop B: $-V_2 + V_3 + V_4 = 0$

Loop C: $-V_4 + V_6 + V_8 = 0$

Loop D: $-V_5 + V_7 - V_6 = 0$

Loop E: $-V_7 + V_9 - V_{10} = 0$

Loop F: $-V_8 + V_{11} = 0$

Loop G: $-V_{11} + V_{10} + V_{12} = 0$

3.6 Parallel and Series Circuits

Two branches are in series if they are connected by a single node and no other branch is connected to that same node. Kirchhoff's current law then states that the current is the same in each branch. A circuit for which all the branches are in series is called a series circuit and the same current passes through each of the branches (see Figure 3.6).

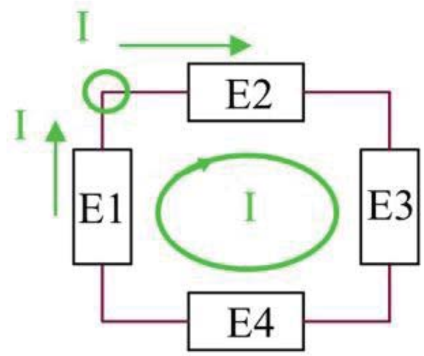


Figure 3.6: Example series circuit.

Two branches are in parallel if they are connected directly to the same two nodes. Kirchhoff's voltage law then states that the voltage across each branch is the same (assuming that they are measured the same polarity).

A circuit for which all branches are connected to the same two nodes is called a parallel circuit (see Figure 3.6).

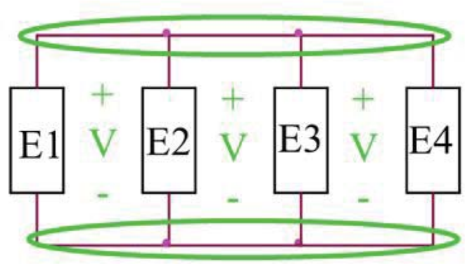
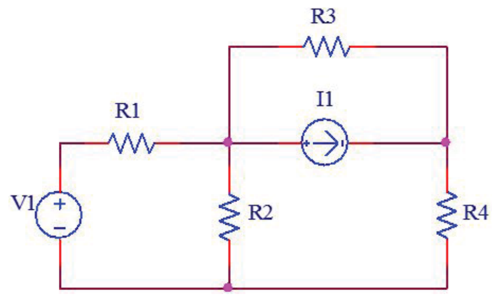


Figure 3.7: Example parallel circuit.

Think About It

Which elements are in series and which are in parallel?



Solution

V_1 and R_1 are in series. R_3 and I_1 are in parallel.

3.7 Measuring Current and Voltage

Consider the circuit show in Figure 3.8(a). We want to measure the current through R4. My shorthand notation for making this measurement is to show the direction of the current measured with an arrow as shown. Measuring current implies that we are measuring the movement of charge. In order to make this measurement the charge needs to pass through the ammeter. Physically, in order to make the current measurement, the circuit must be broken in the path of the desired current and an ammeter put in-line with that path (see Figure 3.8(b)). The polarization of the ammeter must be in passive sign convention with the desired current direction. Any voltage drop across the ammeter will change the current that is being measured! However, if the ammeter voltage is VERY small then the circuit does not change that much. In this class we will assume an ideal ammeter with zero internal resistance, so the voltage drop across the meter is zero. In EE203 you will investigate what happens for real ammeters when the internal resistance is not zero.

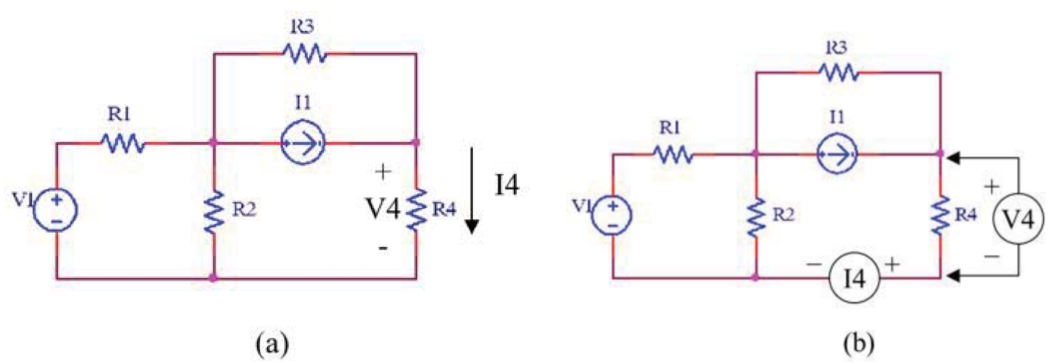


Figure 3.8: Example for measuring current (a) and voltage (b).

We would also like to measure the voltage across R4. My shorthand notation for this measurement is shown in Figure 3.8(a) where the + and − indicates the polarity of the measurement. Remember that voltage is the electric potential difference between these points.

Kirchhoff's Voltage Law tells us that if you want the voltmeter to see the same voltage as that across R4 then we need to put the meter in parallel with R4. Branches that are in parallel have the same voltage across them. So we need to connect the voltmeter across R4 in the same polarity of the desired measurement (see Figure 3.8(b)). Again connecting the voltmeter to the circuit will cause a small amount of current to pass through the meter and not through R4, reducing the voltage across R4. In this class we will assume an ideal voltmeter with infinite internal resistance, so the current passing through the meter will be zero. Again, in EE203 you will investigate the impact of real meters on circuits.

3.8 Circuit Analysis

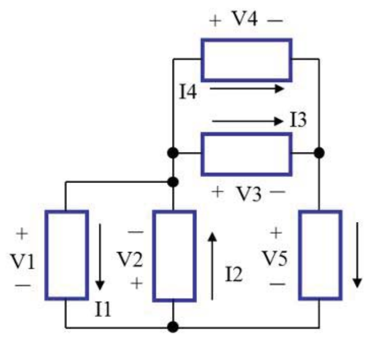


Figure 3.9: Example circuit.

Circuit analysis is the methodical application of the laws that we know to solve for unknown voltages and currents from which all other variables can be derived. The laws that we know so far are:

- Kirchhoff's Current Law: $\Sigma I_{in} = \Sigma I_{out}$
- Kirchhoff's Voltage Law: $\underset{\text{closed loop}}{\Sigma} V_i = 0$
- Conservation of Energy: $\Sigma P_i = 0$

Remember that when calculating power from measured voltages and currents, the measurements must be in passive sign convention with each other.

Consider the circuit shown in Figure 3.9. We would like to determine all currents through and voltages across all elements in the circuit. Where do we start?

1. Arbitrarily choose either the voltage polarity or current directions for the measured voltages or currents.
2. Use passive sign convention to specify the required measured current direction or voltage polarity for the measurement not specified in Step 1.
3. Apply KVL and/or KCL and/or conservation of energy to solve for all unknowns.

Think About It

The following measurements were taken on the circuit shown in Figure 3.9: V_1=2 V, V_4=6 V, I_1=20 mA, I_5=−50 mA, and P_3=60 mW. We would like to determine V_5, I_2, and P_4 and identify which elements are loads and which are sources.

Solution

Writing KVL around the outside loop we have:

$$-V_1 + V_4 + V_5 = 0 \Rightarrow V_5 = V_1 - V_4 = 2 - 6 = -4 \text{ V}$$

Writing KCL we find:

$$I_2 = I_1 + I_5 = 20 - 50 = -30 \text{ mA}$$

To solve for P4 we use what we know about V3:

$$P_3 = V_3 * I_3$$

We need to first calculate I_3 from P_3 (so need to know V_3=V_4) and we find:

$$P_3 = V_3 * I_3 \Rightarrow I_3 = \frac{0.06}{6} = 10 \text{ mA}$$

then we need to calculate I_4 once we have I_3:

$$I_3 - I_4 - I_5 = 0 \Rightarrow 10 - 50 \Rightarrow I_4 = -60 \text{ mA}$$

and then we can calculate P_4.

$$P_4 = I_4 * V_4 \Rightarrow P_4 = -0.06 * 6 = -360 \text{ mW}$$

$$(\text{Source}) \Rightarrow P < 0$$

Finally we need to calculate P for all elements and determine if the P is + or − to determine source or load. (Check your calculations with Table: A)

Table A: Calculated voltage, current and power values for the example circuit.

Branch	**Voltage(V)**	**Current(mA)**	**Power(mW)**
1	2	20	40
2	-2	-30	60
3	6	10	60
4	6	-60	-360
5	-4	-50	200

3.9 Sources

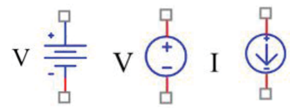
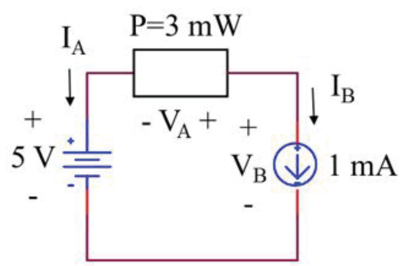


Figure 3.10: Symbols for ideal voltage and current sources.

3.9.1 Ideal Independent Sources

Figure 3.10 shows some symbols for direct current or DC sources. The two on the left are voltage sources and the one on the right is a current source. An ideal voltage source will provide whatever current the circuit requires so as to maintain a constant voltage across its terminals. An ideal current source will provide whatever voltage the circuit requires so as to maintain a constant current passing through it.

Think About It
Consider the following circuit.



Solution
Known:
$I_B = 1$ mA and $I_A = -I_B = -1$ mA
V_A and I_A are in passive sign convention so $V_A = 3$ mW$/I_A$
Apply KVL $(-V - V_A + V_B = 0 \Rightarrow V_B = 5 - 3 = 2$ V$)$
Finally solve for all powers and determine which elements are loads and which are sources.
$P = -5$ mW, $P < 0$, source. $P_A = 3$ mW , $P_B = 2$ mW, $P > 0$ so both are loads.

3.9.2 Ideal Dependent Sources

A dependent source can be either a voltage source or a current source. The output of a dependent source is controlled by a voltage or current which exist in a different part of the circuit. There are four different types of dependent sources (see Figure 3.11): a voltage controlled voltage source, a current controlled voltage source, a current controlled current source, and a voltage controlled current source. For example, the amount of current supplied by the voltage controlled current source shown in Figure 3.11 (d) is I=7V$_B$ where V$_B$ is measured somewhere else in the circuit.

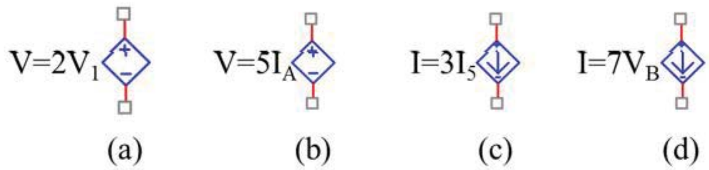
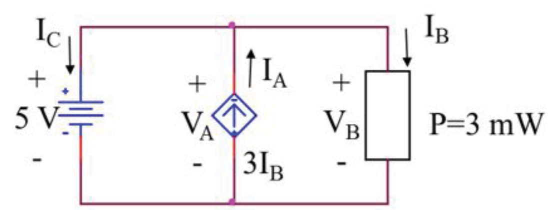


Figure 3.11: Dependent sources:(a) voltage controlled voltage source, (b) current controlled voltage source, (c) current controlled current source, and (d) voltage controlled current source.

Think About It

Consider the following circuit. We would like to determine all currents and voltages in the circuit.



Solution
Since this is a parallel circuit we know that V$_A$ = V$_B$ = 5 V.
We now know I$_B$ = 3 mW/VB
We can calculate I$_C$ because KCL says that I$_A$ = I$_C$ + I$_B$.
We can calculate the power for each element:
P$_C$=5*I$_C$, P$_A$=V$_A$ $*$ $(-$I$_A)$ (needs to be in passive sign convention), and I$_B$$*V_B$ =P=3 mW
Now we can see V$_A$ is a source, and V$_B$ and V are loads.

3.9.3 Real Sources

What happens when you connect a 9 V battery in parallel with a 1.5 V battery as shown in Figure 3.12? Applying Kirchhoff's voltage law around this circuit seems to fail! What is wrong?

It is useful to remind us at this point that the symbols we use are models for real devices, they are not the real devices themselves. If we come across a violation of KVL and/or KCL it is our model for the device that is not appropriate. In this case, the ideal source model for the two batteries is not adequate to determine what happens if the two batteries are connected in parallel. Reviewing section 2.5.1, we see that the chemical reaction inside the battery controls the rate of electrons that can flow between the terminals limiting the maximum current. We model the effect of the chemical reaction ny an *internal resistance* (see Figure 3.13).

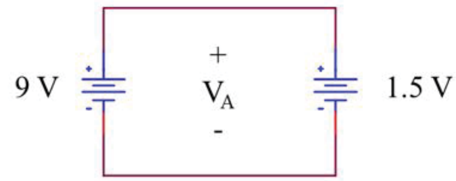


Figure 3.12: Circuit with real sources.

Typically the battery's internal resistance is small and for most circuits could be ignored. However, in this case it cannot.

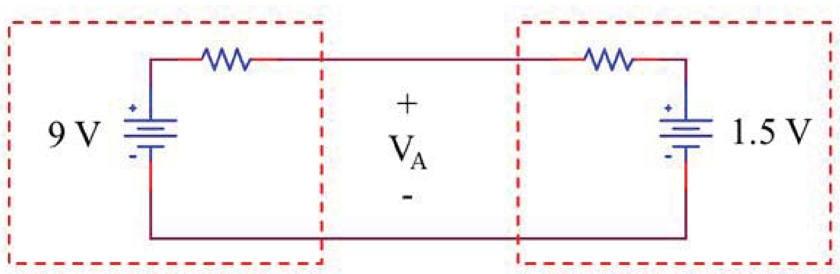


Figure 3.13: Circuit where the real sources are modeled including their internal resistance.

Chapter 4: Resistance

4.1 Key Terms

- **Resistance** can be thought of as the opposition to current.
- **Resistance** of a conductor is given by $R = \rho L/A$, where ρ is the resistivity of the material, L is the length, and A is the cross-sectional area.
- **Ohm's Law** states that the voltage across a conductor is proportional to its resistance and the current passing through it: $V = IR$.

4.2 History

Georg Simon Ohm (1789-1854)

George Ohm discovered in 1827 that current flowing through a conductor is proportional to the voltage across the conductor and inversely proportional to the conductor's resistance.

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

Ohm also experimentally showed that resistance of a conductor depended on the type of material, and its length and cross-section.

$$R = \frac{\rho L}{A}$$

Ohm was ridiculed for his results and was forced to resign his teaching position. Although Georg Ohm discovered one of the most fundamental laws of current

4.3 Characteristics of Resistors

Resistors come in a variety of different types from the very small chip resistors (f) to the very large power resistors (a). Some resistors have a fixed value and some have a value that is variable. Variable resistors may be varied mechanically like the potentiometer (g) or due to other external parameters like temperature (i.e. thermistor (h)), or light (i.e. photoresistor).

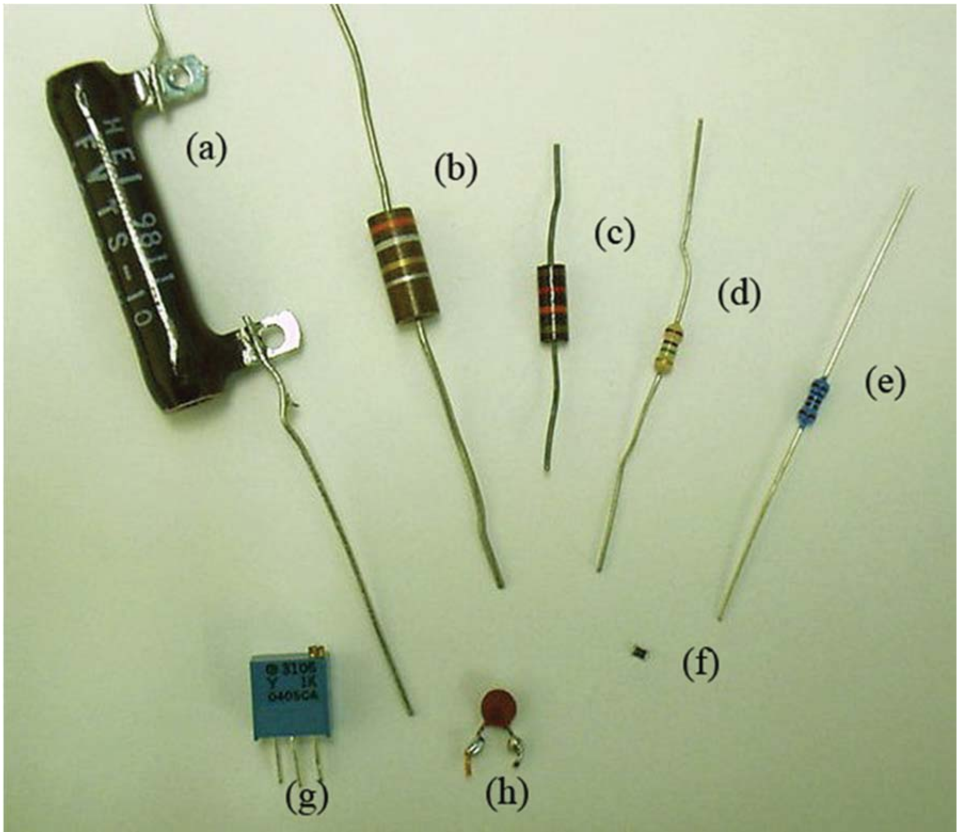


Figure 4.1: Examples of resistors

Resistors are defined by their resistance value, power rating, tolerance, and sometimes temperature coefficient. The power rating of a resistor is often determined by the resistor size. Larger resistors typically can dissipate more power. The resistance value and tolerance are given by colored bands printed directly onto the resistor or by printed numbers as is common for chip resistors.

Table 4.1 gives the definitions for the colors used in the resistor color code. Figure 4.2 gives three examples of how the resistance value could be marked on a resistor. The color bands define the significant digits of the resistance value, the multiplier, and the tolerance. For example, a 120 kΩ 5% resistor would have the color bands: brown = 1, red=2, yellow = x10 kΩ (i.e. 10^4), gold=5%. The tighter tolerance of a 1% resistor requires an additional color band to show an additional significant digit. For example a 274 kΩ 1% resistor has the color bands: red=2, violet=7, yellow=4, orange = x 1 kΩ (i.e. 10^3), brown = 1%. For chip resistors the number markings on the resistor gives the first digit and second digit of the value and the third number gives the times 10 multiplier. So the resistor with the marking 471, is really 47×10^1 or 470 Ω.

Table 4.1: Resistor color code

Color	Digit	Multiplier	Tolerance
Silver		x 0.01 Ω	±10%
Gold		x 0.1 Ω	±5%
Black	0	x 1 Ω	
Brown	1	x 10 Ω	±1%
Red	2	x 100 Ω	±2%
Orange	3	x 1 kΩ	
Yellow	4	x 10 kΩ	
Green	5	x 100 kΩ	±0.5%
Blue	6	x 1 MΩ	±0.25%
Violet	7	x 10 MΩ	±0.1%
Grey	8	x 100 MΩ	
White	9	x 1 GΩ	

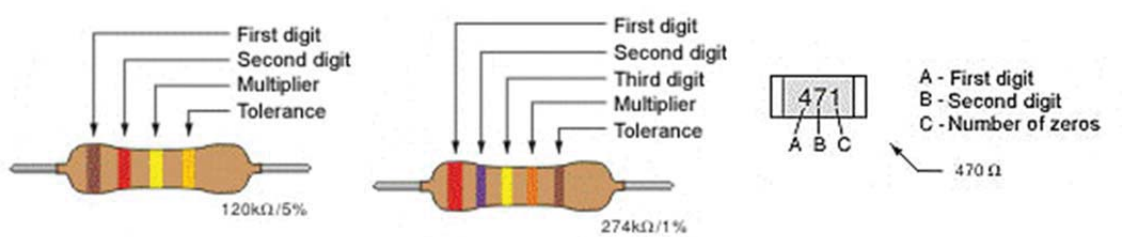


Figure 4.2: Resistor markings

Remember that depending on the atomic structure of a material it may be termed a conductor, semi-conductor or an insulator. These terms relate to how many free electrons are available to move in the material. Conductors have many free electrons. Remember that when a voltage is applied to a conductor the electrons are pushed preferentially in one direction. As they move electrons bump into other electrons and atoms producing heat. The collisions oppose the charge movement and thus limit the current. Hence resistance is defined as the opposition to current, or electron flow.

Think About It

something here.

4.3.1 Resistance of Conductors

The resistance of a conductor is dependent on the type of material. Each material has a certain resistivity given here by the symbol ρ and has the units of Ωm (see Figure 4.3). The resistance of a conductor is proportional to the length of the conductor and inversely proportional to the cross-sectional area. What Ohm discovered experimentally is that the resistance of a conductor (Ω) is the resistivity of the material (Ωm) times the length of the conductor (m) divided by the cross-sectional area (m^2) (see Equation x.x).

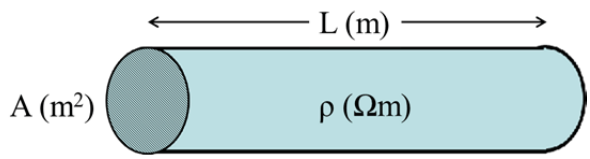


Figure 4.3: Conductor

$$R = \frac{\rho L}{A} \qquad \text{(x.x)}$$

Typical house wiring might be 14 gauge, 12 gauge, or 10 gauge. American wire gauge (AWG) is a standardized wire gauge system used since 1857, mostly in the United States and Canada. Increasing gauge numbers give decreasing wire diameters. So while a 14 gauge wire has a radius of 0.815 mm, a 10 gauge wire has a radius of 1.294 mm. Each wire has a maximum current rating. So a 14 gauge wire would be connected to a 15 A breaker in your breaker box, while a 10 gauge wire would be connected to a 30 A breaker. So why use one size over anther? It depends on what you expect to be connecting to the circuit. 14 AWG is typically used for lighting circuits, 12 AWG for power receptacles, and 10 AWG for heavy current loads such as your oven. The tradeoff is frequently current for cost. A 14 AWG wire contains less copper and therefore costs less than a 10 AWG wire for the same length. So if your circuit doesn't use that much current you can go with the cheaper wire.

Table 4.2 shows the calculation of typical house wire resistance with a length of 75 m using equation x.x. We can see that when the cross-sectional area of the conductor gets larger that the resistance gets smaller. In the 1960's copper was very expensive and contractors turned to using aluminum wire for houses. Again, looking at Table 4.2 we can see that aluminum wire has a larger resistance than copper for the same cross-sectional area and length because the resistivity of aluminum is larger than copper.

Table 4.2: Calculation of typical house wire resistance

L = 75 m	14 AWG (15 A)	12 AGW (20 A)	10 AWG (30 A)
Radius	0.814 mm	1.0265 mm	1.294 mm
$A = \pi r^2$	2.0816×10^{-6} m²	3.3103×10^{-6} m²	5.2604×10^{-6} m²
Copper $\rho = 1.723 \times 10^{-8}$ Ωm	0.621 Ω	0.390 Ω	0.246 Ω
Aluminum $\rho = 2.825 \times 10^{-8}$ Ωm	1.018 Ω	0.640 Ω	0.403 Ω

Think About It

something here.

4.3.2 Temperature Effects

Inherent in all resistors is the variability of their resistance value over temperature. A first order (linear) approximation of the relationship between resistance and temperature is given by equation (x.x), i.e. the change of resistance equals a constant k times the change of temperature.

$$\Delta R = k\Delta T \qquad \text{(x.x)}$$

Resistors that are not thermistors (i.e. temperature dependent resistors) are designed so that their resistance value changes only slightly, or not at all, with changes in temperature. In other words non-thermal resistors have a very small k. A measure of the variability of the resistance over temperature is given by the temperature coefficient, α, of the resistor. Given the resistance value at a given temperature, T_0, and the temperature coefficient, α, for that resistor, the resistance value at a different temperature, T_1, is given by:

$$R_{T_1} = R_{T_0}\left(1+\alpha\left(T_1 - T_0\right)\right) \qquad \text{(x.x)}$$

Examining equations x.x and x.x we see that $\alpha = k / R_{T_0}$. In other words, k, is the *absolute* change of resistance for a given temperature change while α is the *relative* change of resistance, i.e. relative to the resistance value where it is determined. The temperature coefficient may be positive or negative. Figure 4.4 shows the temperature relationship for three different resistors. Curve 9a0 has a large positive temperature coefficient (PTC), i.e. its resistance increases with increasing temperature. Curve (b) has a small positive temperature coefficient and curve (c) has a small negative temperature coefficient. Each of these curves is linear so that *k* is a constant and is equal to the slope of the curve.

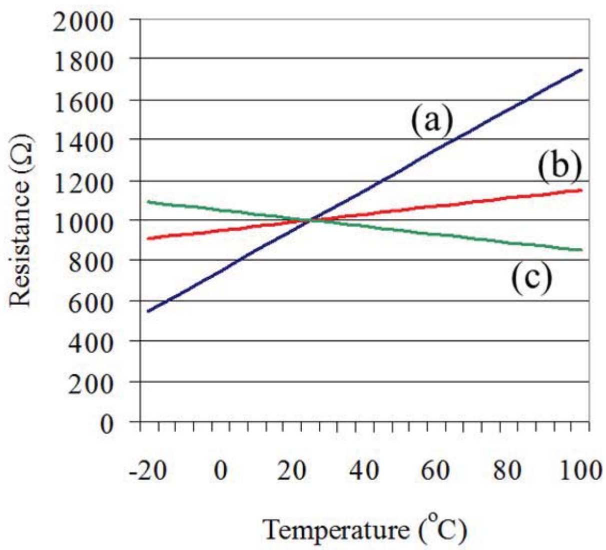


Figure 4.4: Example of three resistors with linear temperature response

Think About It

something here about calculating a resistance given a curve…..

Equations x.x and x.x are good approximations so long as the change of temperature is small. Thermistors are typically used over a wide range of temperatures and do not have a linear relationship between their resistance and temperature. Figure 4.5 shows a family of curve for various thermistors. These curves are modeled using equation x.x. The red curve is represents the calculation of x.x for the following parameters: B = 4100 °K, R_{T0} = 10 kΩ, and T_0 = 25 °C. In equation x.x, exp(x) = e^x, where x must be unitless. Therefore, T_0 and T_1 must first be written in °K.

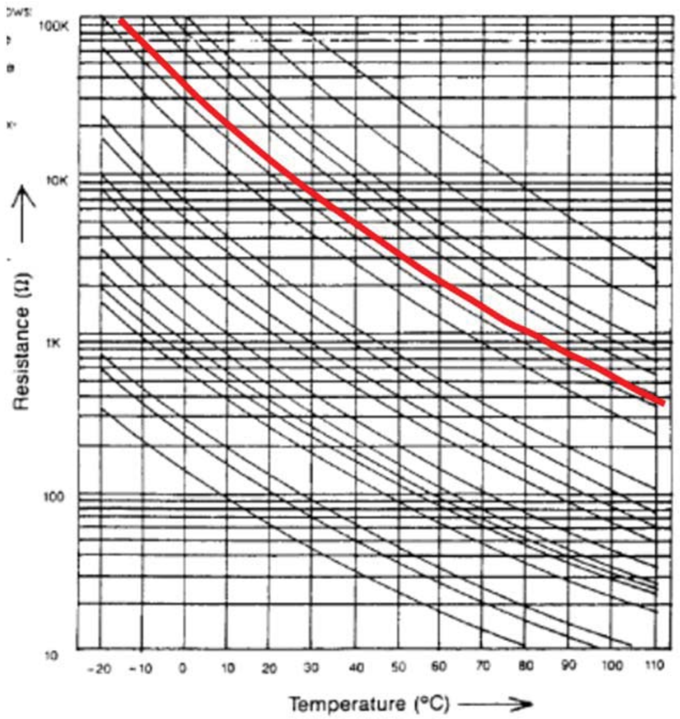


Figure 4.5: Example resistance versus temperature curves for thermistor (reproduced from datasheet).

$$R_{T_1} = R_{T_o} \exp\left[B\left(\frac{1}{T_1} - \frac{1}{T_0} \right) \right] \qquad \text{(x.x)}$$

4.4 Ohm's Law

Ohm's Law states that in any conductor the potential difference (V) across it, in the direction of current flow, bears a constant ratio to the current flow (I) through it.

$$V=IR \qquad \text{(x.x)}$$

Through algebraic manipulation you can write equation x.x as I=V/R or R=V/I depending on which variables you know and which you wish to determine. Remember that the voltage and current must be measured in passive sign convention with each other in order to have the correct sign for the calculated value.

Think About It

something here.

4.4.1 Equivalent Resistance

Series Circuit

Consider the series circuit shown in Figure 4.6. We would like to solve this circuit for all unknown voltages and currents. Writing KVL around the loop gives: $12=V_1+V_2+V_3+V_4+V_5+V_6$. This equation has six unknowns. In order to solve for these unknowns, we need at least six equations. The other necessary equations are obtained by applying Ohm's Law to the voltages of each resistor, i.e. $V_1 = IR_1$, $V_2 = IR_2$, etc. Plugging back into the original KVL equation gives:

$$12 = 3I + 12I + 20I + 15I + 2I + 8I \qquad \text{(x.x)}$$

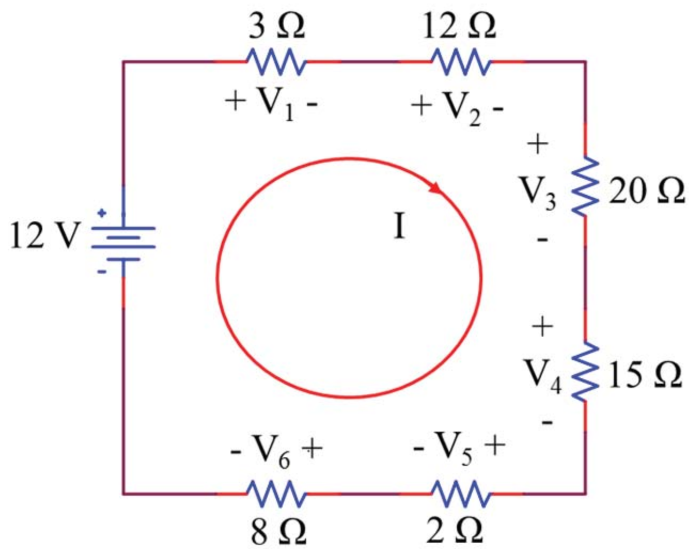


Figure 4.6: Example series circuit.

We can now solve equation x.x for the unknown current:

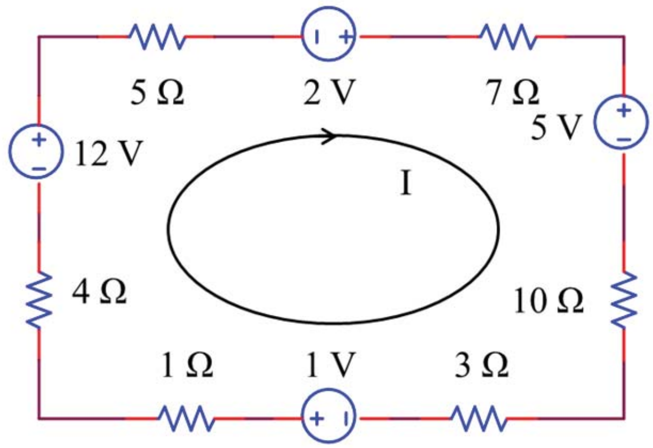
$$I = \frac{12}{(3+12+20+15+2+8)} = \frac{V_{total}}{R_{eq}} \qquad \text{(x.x)}$$

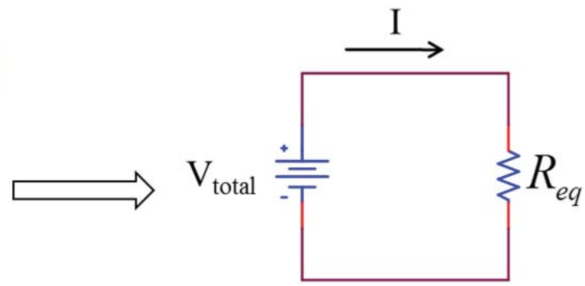
Equation x.x looks like Ohm's Law where we have replaced the six individual resistors with an equivalent series resistor. The equivalent series resistance of a circuit or part of a circuit with *N* resistors in series is given by:

$$R_{eq} = \sum_{i=1}^{N} R_i \qquad \text{(x.x)}$$

Think About It

What is the total series voltage and the total series equivalent resistance?





Parallel Circuit

Consider the parallel circuit shown in Figure 4.7. We would like to solve this circuit for all unknown voltages and currents. Writing KCL at the top node gives: $10=I_1+I_2+I_3+I_4+I_5$. This equation has five unknowns. In order to solve for these unknowns, we need at least five equations. The other necessary equations are obtained by applying Ohm's Law to the currents of each resistor, i.e. $I_1 = V/R_1$, $I_2 = V/R_2$, etc. Plugging back into the original KCL equation gives:

$$10\,[\text{mA}] = \frac{V}{2}+\frac{V}{5}+\frac{V}{10}+\frac{V}{7}+\frac{V}{11} \qquad \text{(x.x)}$$

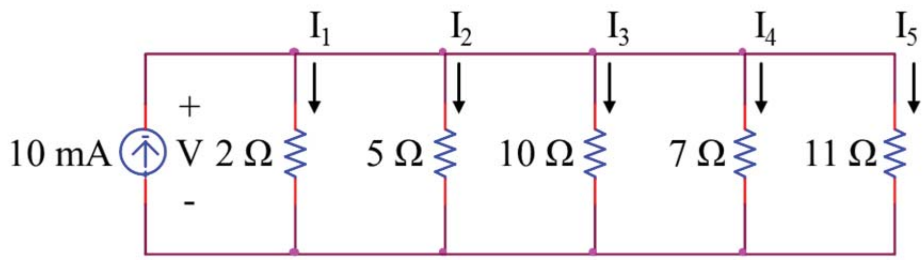


Figure 4.7: Example parallel circuit

We can now solve equation x.x for the unknown voltage:

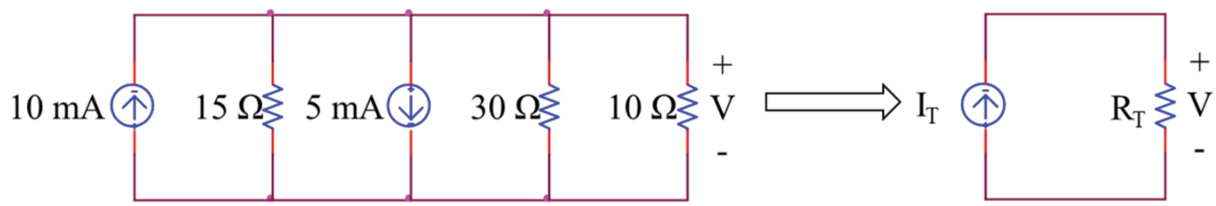
$$V = (10)\frac{1}{\left(\frac{1}{2}+\frac{1}{5}+\frac{1}{10}+\frac{1}{7}+\frac{1}{11}\right)} = I_{total}R_{eq} \qquad \text{(x.x)}$$

Equation x.x looks like Ohm's Law where we have replaced the five individual resistors with an equivalent parallel resistor. The equivalent parallel resistance of a circuit or part of a circuit with *N* resistors in parallel is given by (one over the sum of the one over's):

$$R_{eq} = \frac{1}{\sum_{i=1}^{N} \frac{1}{R_i}} \qquad \text{(x.x)}$$

Think About It

What is the total parallel current and the total parallel equivalent resistance?



4.4.2 Resistor Circuits

Consider the circuit shown in Figure 4.8 (a). We would like to determine all voltages and currents in this circuit. On approach to solving this circuit is to determine the equivalent resistance seen by the 10 V source (see Figure 4.8 (b-c)) and then solve for I_1. Once I_1 is determined the remaining unknowns are derived through the application of KVL, KCL, and Ohm's Law.

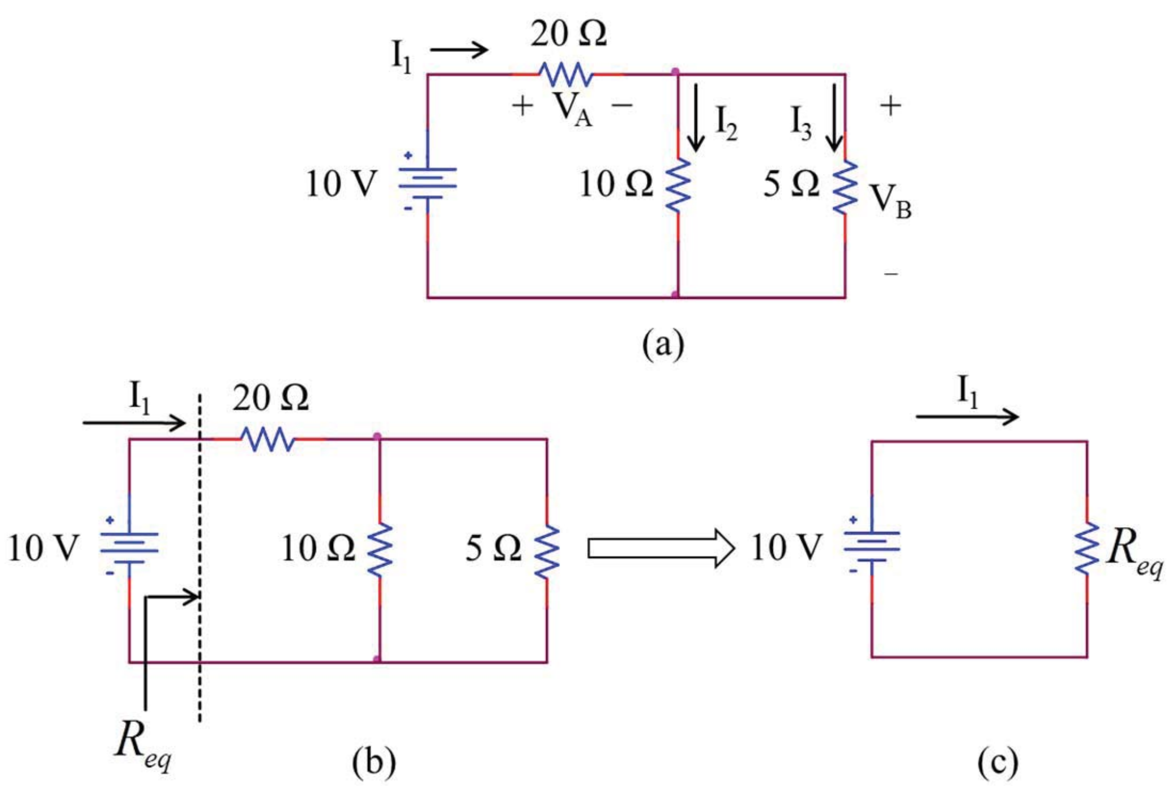


Figure 4.8: Example resistor circuit

To determine the circuit equivalent resistance we first recognize that the 10 Ω and 5 Ω resistors are in parallel, so we apply equation x.x to determine the equivalent parallel resistance of those two resistors. This equivalent parallel resistance is in series with the 20 Ω resistor so we apply equation x.x to determine the circuit equivalent resistance. Combining all this steps into a single equations gives:

$$R_{eq} = 20 + \frac{1}{\frac{1}{10}+\frac{1}{5}} = 23.\overline{3}\,\Omega \qquad \text{(x.x)}$$

Once we have the circuit equivalent resistance, I_1 is determined by using Ohm's Law:

$$I_1 = \frac{V}{R_{eq}} = \frac{10}{23.\overline{3}} = 0.42857\text{ A} \qquad \text{(x.x)}$$

If we then write KVL around the first loop we get $-10+V_A+V_B=0$. $V_A = 20I_1$, and we have already solved for I_1. Therefore:

$$V_B = 10 - 20I_1 = 1.42857\text{ V} \qquad \text{(x.x)}$$

We recognize that VB is the voltage across both the 10 Ω and 5 Ω resistors, so applying Ohm's Law we determine:

$$\begin{aligned} I_2 &= \frac{1.42857}{10} = 0.14286\text{ A} \\ I_3 &= \frac{1.42857}{5} = 0.28571\text{ A} \end{aligned} \qquad \text{(x.x)}$$

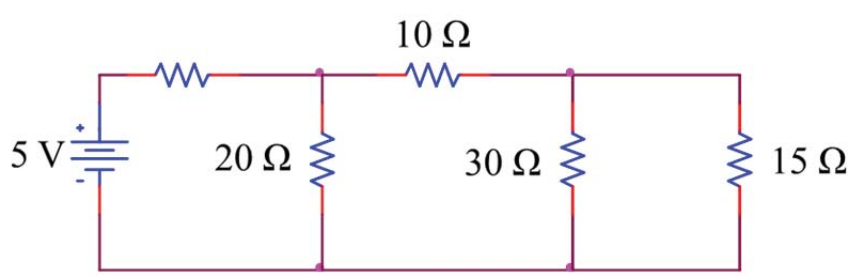
We can check our solution by applying KCL to the top node:

$$I_1 = I_2 + I_3 = 0.14286 + 0.28572 = 0.42857\text{ A} \qquad \text{(x.x)}$$

Equations x.x and x.x give the same result, so our solution checks.

Think About It

Solve for the equivalent resistance seen by the 5 V supply, then solve for other voltages and currents.



Chapter 5

Circuit Analysis

In this chapter we bring together the concepts of the previous chapters and develop a systematic method of circuit analysis that may be applied to any arbitrary circuit. The method introduced in this chapter, Nodal Analysis, is one of several circuit analysis methods. Alternative methods will be discussed in EE203. As a reminder, the four main laws used in circuit analysis are repeated here.

- Kirchhoff's Current Law: $\sum I_{in} = \sum I_{out}$
- Kirchoff's Voltage Law: $\sum\limits_{\text{closed loop}} V_i = 0$
- Conservation of Energy: $\sum P_i = 0$
- Ohm's Law: V=IR

5.1 Differential versus Node Voltages

The terms differential and node voltages were introduced in Chapter 2. We will now expand our discussion to include the use of node voltages when applying Ohm's Law. First recall the difference between differential and node voltages.

Consider the circuit shown in Figure 5.1. The labels A, B, C, and D identify all the nodes in the circuit. Remember that a node is the point of connection between two or more branches. Voltage is defined as the potential difference between any two points or nodes. Specific nodes can be identified for a measurement by using subscripts on the voltage variables. For instance V_{AB} is defined as the *differential voltage* measured at node A relative to node B, i.e. the positive terminal of the voltmeter is connected to node A and the negative terminal of the voltmeter is connected to node B. V_{BC} is the voltage measured at node B relative to node C, V_{BD} is the voltage measured at node B relative to node D. These measurements are called *differential voltages* to emphasize that voltage is measured between two nodes.

A *node voltage* is that voltage measured between an arbitrary node and a specific reference node, sometimes called ground. For example, in the circuit shown in Figure 5.1, node D might be defined as the reference node or ground. Then the *node voltage* measured from node B to the reference node, node D is defined as V_B. As you can see from the circuit, $V_B = V_{BD}$. We can also define the *node voltages* for node A, V_A, $V_A = V_{AD}$, and node C, $V_C = V_{CD}$. V_{AB} and V_{BC} are not *node voltages* since they are not measured relative to our pre-defined reference node D.

It is frequently the case that only node voltages are measured. We would like to know if there is a way to determine differential voltages form our node voltage measurements. We can write Kirchhoff's voltage law around the left hand loop:$-V_A + V_{AB} + V_B = 0$. Solving this equation for V_{AB} we get:

$$V_{AB} = V_A - V_B \tag{5.1}$$

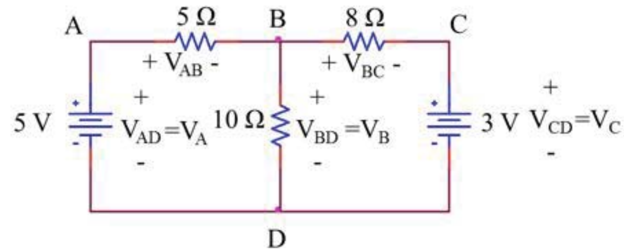


Figure 5.1: Example circuit comparing differential and node voltages.

In other words, the differential voltage V$_{AB}$ equals the difference of the node voltages, V$_A$ and V$_B$ on both sides of the measurement. Similarly the differential voltage V$_{BC}$ is given by by: $V_{BC} = V_B - V_C$, where the node voltage, V$_C$, attached to the negative terminal of the V$_{BC}$ measurement is subtracted form the node voltage, V$_B$, measured at the positive terminal of the VBC measurement.

Consider the partial circuit shown in Figure 5.2.

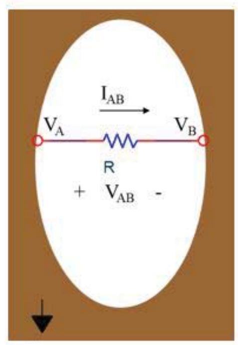


Figure 5.2: Example partial circuit.

When a circuit is drawn this way you can imagine that you are looking at a portion of a larger circuit, perhaps through a window. You have access to the node voltages V$_A$ and V$_B$ that have been measured relative to some reference in the circuit, but you do not know what created those voltages. We are interested in determining V$_{AB}$ and I$_{AB}$. Using Ohm's Law we can determine I$_{AB}$ from V$_{AB}$.

$$I_{AB} = \frac{V_{AB}}{R} = \frac{V_A - V_B}{R} \tag{5.2}$$

Examining Equation 5.2 we remind ourselves that Ohm's Law states that the current through a resistor is equal to the voltage across the resistor divided by the resistance. Remember that the voltage across the resistor is **always** a *differential voltage*! We use the node voltages on either side of the resister to calculate the differential voltage across the resistor.

Consider the partial circuit shown in Figure 5.3. We know the node voltages V$_B$, V$_C$, and V$_D$, we want to determine V$_A$, I$_1$, I$_2$, and I$_3$. As with all circuit analysis we apply KCL, KVL and Ohm's Law to determine unknown parameters.

We can write KCL at node A:

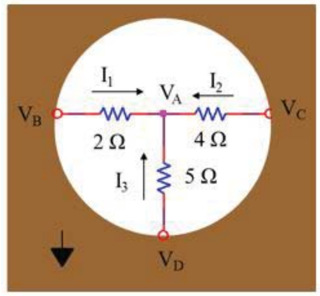


Figure 5.3: Example partial circuit.

$$I_1 + I_2 + I_3 = 0 \tag{5.3}$$

We have three unknowns and only one equation. We need additional equations. We can convert the currents in Equation 5.3 using Ohm's Law to voltage across the resistor divided by the resistor.

$$\frac{V_{AB}}{2} + \frac{V_{AC}}{4} + \frac{V_{DA}}{5} = 0 \tag{5.4}$$

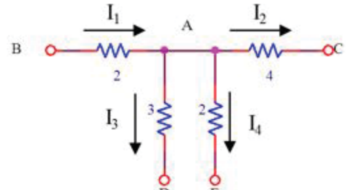
We need to be careful to use the voltage across the resistor measured in passive sign convention with the measured direction of the current. For example, I_3 is measured going from node D to node A. Therefore the voltage that is in passive sign convention with I_3 is V_{AD}. Equation 5.4 still has three unknowns, however if I write the differential voltages in terms of node voltages Equation 5.4 reduces to:

$$\frac{V_B - V_A}{2} + \frac{V_C - V_A}{4} + \frac{V_D - V_A}{5} = 0 \tag{5.5}$$

In Equation 5.5 only V_A is unknown. Using this equation we can solve for V_A, then determine all the differential voltages V_{BA}, V_{CA}, V_{DA} and currents I_1, I_2, and I_3.

Think About It

Determine the node voltage V_A and the current I_2, given that V_B=5 V, $V_C = -5$ V, V_D =10 V, and V_E = -20 V.



Solutions

$$I_1 = I_2 + I_3 + I_4$$

$$\frac{V_{AB}}{2} = \frac{V_{AC}}{4} + \frac{V_{AD}}{3} + \frac{V_{AE}}{2}$$

$$\frac{V_B - V_A}{2} = \frac{V_A - V_C}{4} + \frac{V_A - V_D}{3} + \frac{V_A - V_E}{2}$$

$$\frac{5-V_A}{2} = \frac{V_A-(-5)}{4} + \frac{V_A-10}{3} + \frac{V_A-(-20)}{2} \Rightarrow V_A = -3.42 \text{ V}$$

$$I_{@} = \frac{V_{AC}}{4} = \frac{V_A - V_C}{4} = \frac{(-3.42)-(-5)}{4} = 0.395 \text{ A}$$

5.2 Nodal Analysis

In circuit analysis we are trying to find the current through each element and the voltage across each element. Analysis implies that the circuit elements have known values. In other words, analysis implies that we know the value of each resistor and source.

Nodal Analysis outlines a methodology which defines a specific set of steps to perform when doing circuit analysis:

1. Obtain a set of "n" equations for the node voltages.
   - Specify one node as the reference.
   - Write KCL for each of the other nodes.
   - Replace all known current values (i.e. current sources).
   - Use Ohm's Law and KVL to express the remaining currents in terms of node voltages.
2. Solve the obtained "n" equations to find the node voltages at each node.
   - Compute the voltage across each component using KVL.
3. Compute the current through each resistance using Ohm's Law.
4. Compute the remaining currents using KCL.

Let us apply nodal analysis to the partial circuit shown in Figure 5.4. for the time being we will assume we know the node voltages at node B, C, and D.

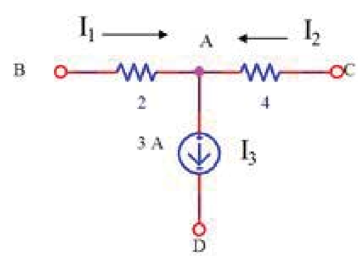


Figure 5.4: Example partial circuit.

Step 1 says that we need to write Kirchhoff's current law at node A. This gives I_1+I_2=3. Note that I have already plugged in for the know current source. Now we apply Ohm's Law to convert all unknown currents into voltages divided by resistance.

$$\frac{V_B - V_A}{2} + \frac{V_C - V_A}{4} = 3 \tag{5.6}$$

Since V_A and V_C are known, Equation 5.6 can be used to solve for V_A.

Consider the partial circuit shown in Figure 5.5 (a), where V_B, V_C, and V_D are known. Writing KCL at node A gives $I_1 + I_2 = I_3$.

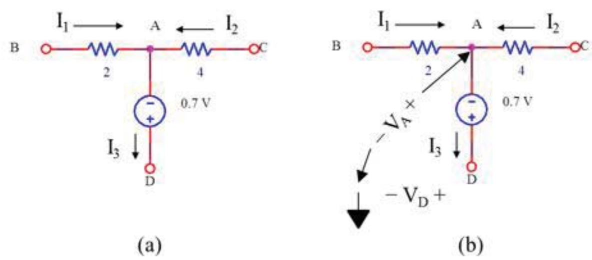


Figure 5.5: Example partial circuit.

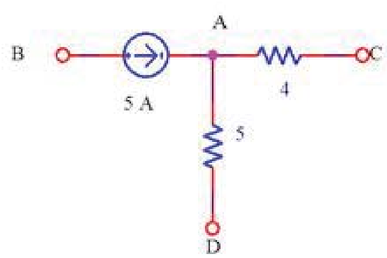
Applying Ohm's Law to each resistor gives:

$$\frac{V_B - V_A}{2} + \frac{V_C - V_A}{4} = I_3 \tag{5.7}$$

I can not substitute in directly for I_3 since the source is not a current source. I can't apply Ohm's Law since it is not a resistor. I need to handle voltage sources that are not connected to our reference node differently. I need to apply KVL around the loop. In this case I can imagine a reference location and write KVL as $-V_D + 0.7 + V_A = 0$ (see Figure 5.6 (b)). So $V_A = V_D - 0.7$. Since we now know V_A, we can solve for I_1 and I_2 and then use Equation5.7 to solve for I_3.

Think About It

Use Nodal Analysis to solve for V_A given that V_B = -2 V, V_C = 10 V and V_D = 7.5 V.



Solution

$$I_1 = I_2 + I_3$$

$$(20)\left[I_1 = \frac{V_A}{5} + \frac{V_A}{4}\right] \Rightarrow 20I_1 = 9V_A$$

$$V_A = \frac{(20)(5)}{9} = 11.11 \text{ V}$$

5.3 Power Dissipation and RMS

Once all voltages and currents are known in a circuit you can assess how much power is being dissipated in each element. This is useful if you want to make sure you don't require any particular element to dissipate more power than it is capable of. Remember that P=VI. For resistors we also know that V=IR. Therefore a resistor:

$$P_R = V_R I_R = I_R^2 R = \frac{V_R^2}{R} \tag{5.8}$$

This equation works great for DC sources, but what if the source was time-varying? Consider the circuit shown in Figure 5.6 where the circuit is driven with a sinusoidal voltage source with amplitude 10 V. We would like to know what the average power dissipation is for both resistors.

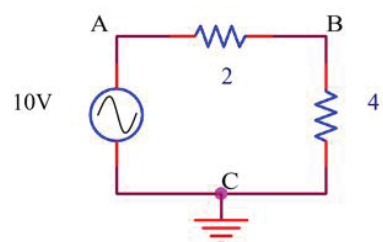


Figure 5.6: Resistor circuit driven by a sinusoidal source.

In order to determine the power dissipation in the resistor we apply Equation 5.8. However, the source is time varying so the equation becomes:

$$P(t) = \frac{V^2(t)}{R} \tag{5.9}$$

Equation 5.9 represents the instantaneous dissipation in the resistor. If the source is periodic (i.e. household AC power) it is meaningful to talk about the "average" power dissipated over time calculated in the same fashion as we calculated the time average of any waveform.

$$P_{ave} = \frac{1}{T}\int_0^T P(t)dt = \frac{1}{T}\int_0^T \frac{V^2(t)}{R}dt = \frac{\frac{1}{T}\int_0^T V^2(t)dt}{R} \tag{5.10}$$

In other words, to find the average power I would integrate the instantaneous power over one period. Since R is not time varying (i.e. R is always a constant), R can be removed from the integration. The numerator in the last term of Equation 5.10 is defined as the square of V_{RMS}.

$$V_{RMS} = \sqrt{\frac{1}{T}\int_0^T V^2(t)dt} \tag{5.11}$$

The RMS (Root-Mean-Squared) value is an equivalent DC value: it tells you how many volts or amps of DC that a time-varying waveform is equal to in terms of its ability to produce average power.

$$P_{ave} = \frac{V_{RMS}^2}{R} = I_{RMS}^2 R \tag{5.12}$$

In this class we restrict our discussion to two different waveforms, sinusoidal and piece-wise constant. In order to determine the RMS value of any waveform we need to first *Square* the waveform, take its *Average*, then take the square-*Root* of the average (Root-Mean-Squared).

$$RMS = \sqrt{ave(x^2(t))} \tag{5.13}$$

Reviewing Chapter 1, we remind ourselves that a generic sinusoid (shown in Figure 5.7) with a potential DC offset is given by the equation:

$$V(t) = A + B\sin(\omega t + \phi) \tag{5.14}$$

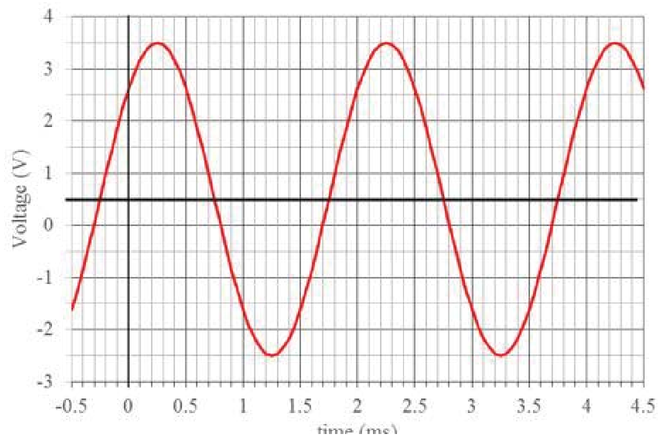


Figure 5.7: $V(t) = 0.5 + 3\sin(\omega t + \phi)$

The square of this waveform (shown in Figure 5.8) is:

$$V^2(t) = A^2 + AB\sin(\omega t + \phi) + B^2\sin^2(\omega t + \phi) \tag{5.15}$$

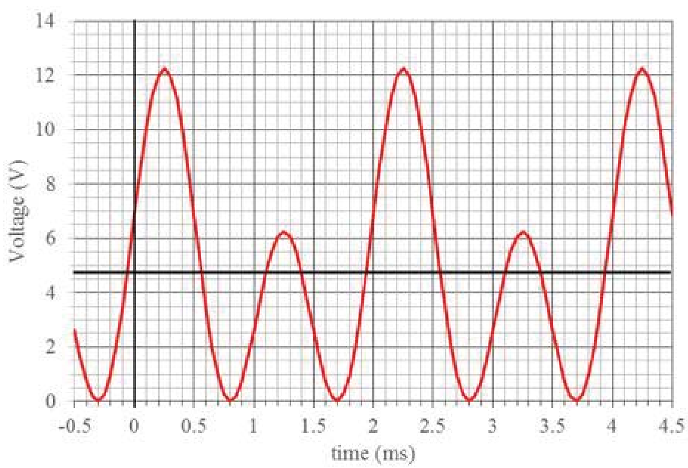


Figure 5.8: $V^2(t) = (0.5)^2 + (0.5)(3)\sin(\omega t + \phi) + (3)^2\sin^2(\omega t + \phi)$

The average of a summation is the sum of the average of each of its terms. So the first term A^2 is a constant. The time average of a constant is just the constant, so the time average of A^2 is A^2. The second term is a sinusoid with an amplitude of AB. The average of a sinusoid is always zero. The last term is the square of a sinusoid multiplied by B^2. The average of the square of a sinusoid is $\frac{1}{2}$. Therefore, the average of $B^2\sin(\omega t + \phi)$ is $B^2/2$.

$$V_{ave}^2 = A^2 + 0 + \frac{B^2}{2} \tag{5.16}$$

The dark horizontal line in Figure 5.8 shows the time average of the waveform, $0.5^2 + \frac{3^2}{2} = 4.75$. Finally, the RMS voltage represents Equation 5.17 is ($V_{RMS} = 2.18$ for our example).

$$V_{RMS} = \sqrt{A^2 + \frac{B^2}{2}} \tag{5.17}$$

Note that if the amplitude B=0 such that V(t) = A, i.e. a DC voltage, then V_{RMS} = A. If the DC offset A=0, so that the original waveform is a pure sinusoid, $V(t) = B\sin(\omega t + \phi)$, then $V_{RMS} = 0.707$ V.

The same process is applied to determine the RMS value of a piece-wise constant waveform. Consider the voltage waveform shown in Figure 5.9. This is the same waveform illustrated in Chapter 1. In order to determine the RMS value we need to first square the waveform. The Square of the waveform is shown in Figure 5.10. Note that for a piece-wise constant waveform, the square if found by taking the square of each individual voltage level. The time average of the waveform shown in Figure 5.10 is obtained in the same way as described in Chapter 1.

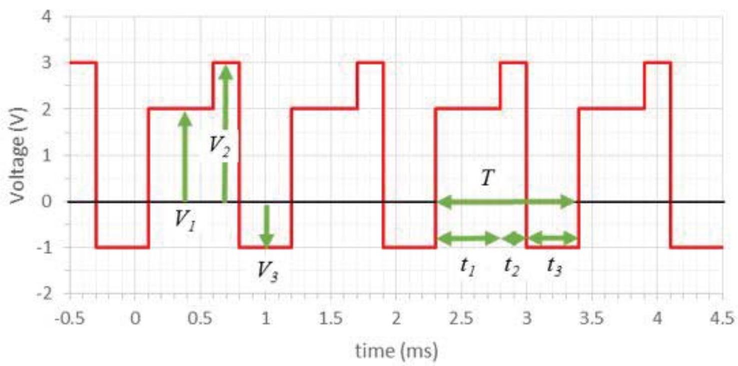


Figure 5.9: V(t)

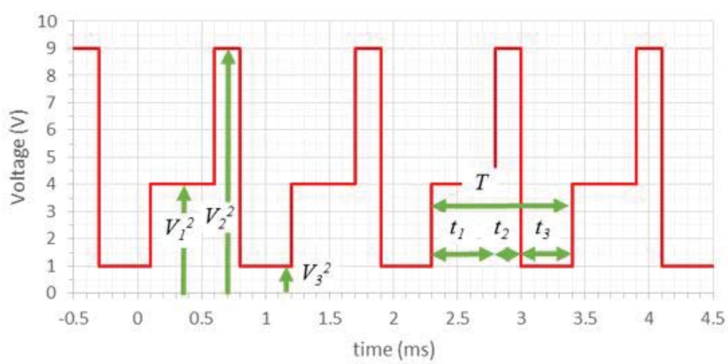


Figure 5.10: $V^2(t)$

The time average of the waveform shown in Figure 5.10 is obtained in the same way as described in Chapter 1.

$$V_{ave} = \frac{\sum_i V_i t_i}{T} \tag{5.18}$$

For our example we have three voltage levels, so the average of the square of the voltage waveform is:

$$V_{ave}^2 = \frac{V_1^2 \cdot t_1 + V_2^2 \cdot t_2 + V_3^2 \cdot t}{T} = \frac{(2\ \mathrm{V}^2)(0.5\ \mathrm{ms}) + (3V)^2(0.2\ \mathrm{ms} + (-1\mathrm{V}^2)(0.4\ \mathrm{ms})}{1.1\ \mathrm{ms}} = 3.82\ \mathrm{V} \tag{5.19}$$

Therefore the RMS value of this piece-wave constant waveform is $\mathrm{V}_{RMS} = 1.95$ V.

Chapter 6: Magnetism and Magnetic Circuits

Many devices rely on magnetism: computer disk drives, VCR's, motors, generators, transformers, speakers, the strip on your credit cards and polar express cards. To understand these devices we need to understand how magnetism works, the relationship between electrical and magnetic quantities, and magnetic circuits.

In this chapter we will cover magnetic field, **H**, magnetic flux, **Φ**, magnetic flux density, **B**, and their relationship in magnetic circuits. We will illustrate a method of determining the direction of the magnetic flux using the right hand rule. We introduce magnetomotive force, $\Im$, and reluctance, $\Re$, and see how those terms, with magnetic flux, are used to describe a magnetic circuit in the same way electric circuits are described in terms of voltage, resistance, and current. Next we examine how a magnetic field is defined in terms of the magnetomotive force and also related to the magnetic flux density through magnetization curves. Finally we will conclude our discussion of magnetism with magnetic forces and magnetic machines.

6.1 Key Concepts

- Magnetic Field, **H**, is the magnetic influence of electric currents and magnetic materials.
- Magnetic Flux, **Φ**, amount of magnetic field that is passing through a surface.
- Magnetic Flux Density, **B**, is the amount of magnetic flux through a unit area taken perpendicular to the direction of the magnetic flux.
- Reluctance, $\Re$, is the resistance to magnetic flux.
- Magnetomotive Force, $\Im$, represents the potential that a hypothetical magnetic charge would gain by completing a closed loop.
- Ampere's Circuital Law: $\Im = NI = \sum_i H_i L_i$
- Magnetic Force: $F = q\vec{v} \times \vec{B}$

6.2 History

Hans Christian Øersted 1777-1851

Hans Christian Øersted, one of the leading scientists of the nineteenth century, played a crucial role in understanding electromagnetism. In 1820 he discovered that a compass needle deflects from magnetic north when an electric current is switched on or off in a nearby wire.

He went on to show that ac conductor carrying a current will be deflected by a magnetic field. This showed that electricity and magnetism were related phenomena, a finding that laid the foundation for the theory of electromagnetism and for the research that later created such technologies as radio, television, and fiber optics.

The unit of magnetic field strength was named the Øersted in his honor.

André-Marie Ampère 1775-1836

On September 11, 1820, Ampère heard of Øersted's discovery that a magnetic needle is acted on by a voltaic current. A week later, on September 18th, Ampère presented a paper to the Academy, containing a far more complete explanation of Øersted's findings. Ampère provided a mathematical formulation for Øersted's discovery.

He also established Ampère's Circuital Law which is to magnetic circuits what Kirchhoff's Voltage Law is to electric circuits. Ampère's Circuital Law states that the sum of the magnetomotive forces around a magnetic circuit (closed loop) equals the sum of the magnetomotive drops around the circuit. Ampère's circuital law is to magnetic circuits

Ampère's theories established the relationship between electricity and magnetism and became the basis of the science of electromagnetism

Michael Faraday 1791-1867

Michael Faraday was one of the greatest experimenters ever. Because he was self-trained, however, he had no grasp of mathematics and was unable to understand any of Ampère papers.

Through the course of his experiments, Faraday discovered that a suspended magnet revolves around a current bearing wire, leading him to propose that magnetism is a circular force. Experimentally, he makes the first conversion of electrical into mechanical energy (electric motor).

Faraday made many other discoveries throughout his life, some of the more notable include: the concept of lines of force, electromagnetic induction (transformers), and two laws of electrochemistry.

James Clerk Maxwell 1831-1879

One of Maxwell's most important achievements was his extension and mathematical formulation of Michael Faraday's theories of electricity and magnetic lines of force. He published his work in the paper "On Faraday's Lines of Force", which was read to the Cambridge Philosophical Society in two parts, 1855 and 1856. In this work, Maxwell showed that a few simple mathematical equations could express the behavior of electric and magnetic fields and their interrelations.

Maxwell is widely regarded as the nineteenth century scientist who had the greatest influence on twentieth century physics. Maxwell spent the majority of his time studying electricity and magnetism. He combined the findings of Faraday, Ampère, Øersted, and others into a linked set of differential equations; these equations are collectively known as Maxwell's Equations. His equations, first presented to the Royal Society in 1864, describe the behavior of both the electric and magnetic fields, as well as their interactions with matter.

6.3 Magnetic Field and Magnetic Flux

We are all familiar with a bar magnet. It has two poles, north and south. If a tray with iron filings is placed on top of a bar magnet, the iron filings will line up in the manner of the lines shown in Figure 6.1. These iron filings give a visual representation of what Faraday called magnetic fields or lines of force.

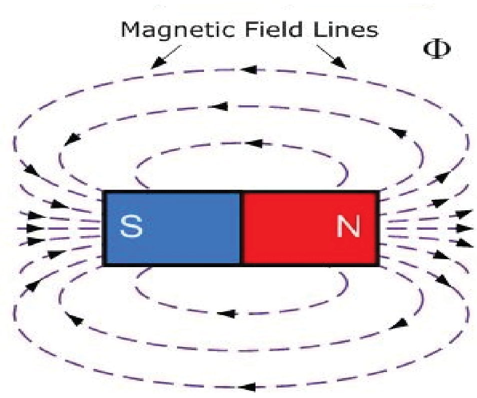


Figure 6.1: Bar magnetic showing magnetic field lines

These lines show the direction of the force that would occur if an isolated north pole was placed anywhere in the field. Faraday's lines of force are called magnetic flux lines. They are represented mathematically by the capital Phi (**Φ**) and have units of Webers (Wb). Magnetic flux lines show the direction and intensity of the magnetic field at all points.

Using the idea of magnetic flux lines we can see why unlike poles attract and like poles repel. Figure 6.3 (a) is an example of how unlike poles' magnetic flux lines behave. Notice that the magnetic flux lines are passing from one magnet to the other; they combine and act to pull the two magnets together.

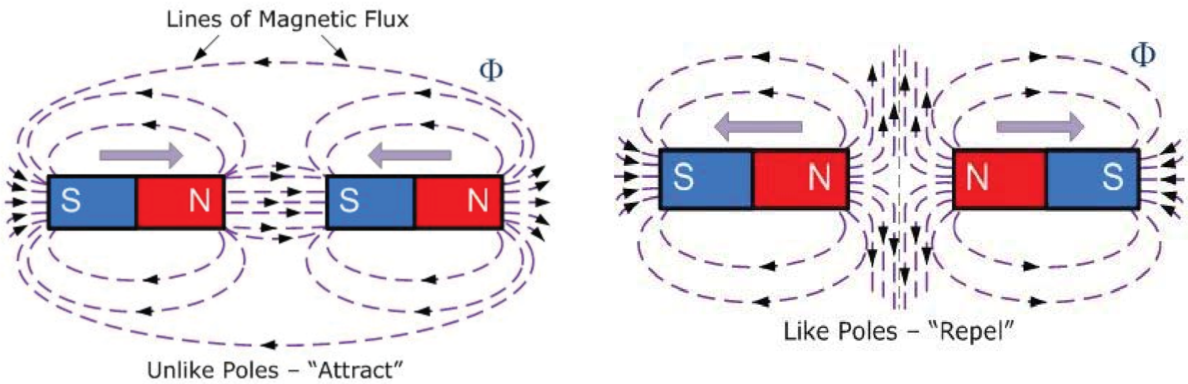


Figure 6.3: Illustration of the (a) attractive behavior between opposite poled bar magnets and (b) repulsive behavior between like poled bar magnets; emphasized here with magnetic field lines drawn

In Figure 6.3 (b) however, the two north poles are placed next to each other. In this case the flux lines from the two magnets oppose each other and cause the two magnets to separate.

Ferromagnetic materials (iron, nickel, cobalt, etc.) provide an easy path for magnetic flux (low reluctance). Air does not (high reluctance). Given a choice of passing through ferromagnetic material or non-ferromagnetic material, magnetic flux lines will go out of their way to pass through the ferromagnetic material, as shown in Figure 6.2. If you provide a complete circuit (closed loop) of ferromagnetic material, then all the magnetic flux will be contained in the material. No flux will be observed outside of the material.

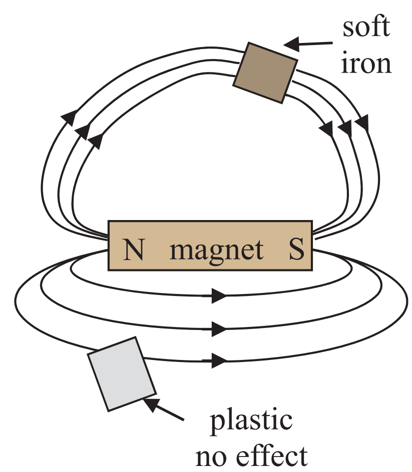


Figure 6.2: Change of magnetic field in the presence of ferromagnetic material

6.4 Magnetic Flux and Flux Density

Magnetic flux, **Φ**, can be thought of as the magnetic current flowing through the material. It is not a current of magnetic charge; it merely has the same relationship in equations describing magnetic circuits. Just like current, magnetic flux has a direction associated with it. Magnetic flux density, **B**, is the distribution of magnetic flux that passes through a specified area. Since the magnetic flux might pass through the area going either way, the magnetic flux density also has a direction associated with it. For magnetic flux that passes perpendicularly through a surface, we define the magnetic flux density on a surface as the total flux divided by the area of the surface, as shown in Figure 6.4 ($\mathbf{B_1}$)

$$\vec{B} = \frac{\Phi}{A}\hat{n} \tag{x.x}$$

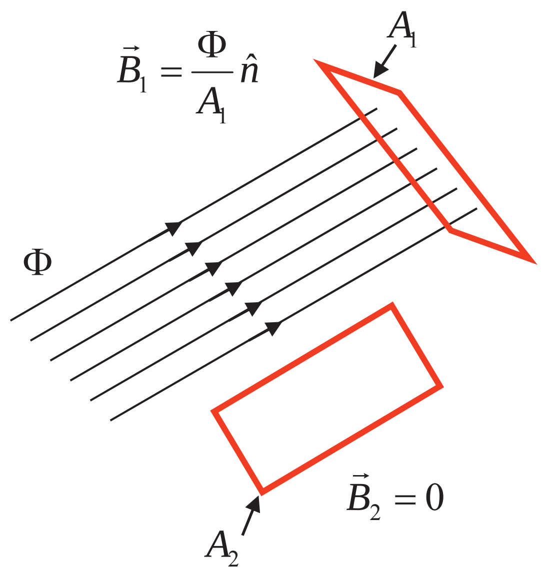
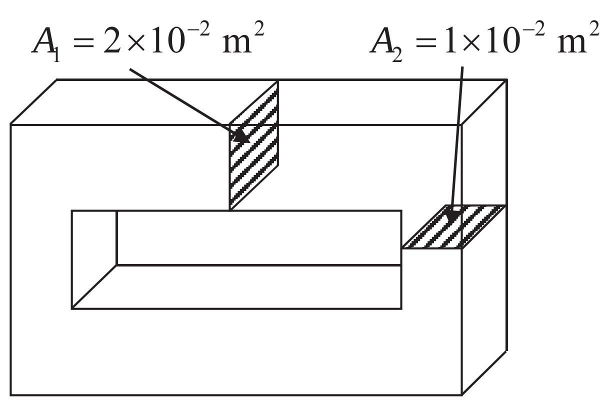


Figure 6.4: Illustration showing the measurement of magnetic flux density on a surface.

The arrow on top of the symbol for the flux density indicates that this quantity is a vector, i.e. has a direction. The $\hat{n}$ indicates that the measured direction of the magnetic flux density is in the unit normal direction, i.e. perpendicular, to the surface area. Figure 6.4 shows two measurements for magnetic flux density, $\mathbf{B_1}$ measured on A_1 and $\mathbf{B_2}$ measured on A_2. $\mathbf{B_1}$ is the total flux divided by the area in the $\hat{n}$ direction. This is because the flux direction is perpendicular to the area. $\mathbf{B_2}$, however, is zero, since the area is parallel to the flux lines and therefore no flux passes through the area. In this class, we simplify our equations by only considering geometries for which the flux is perpendicular to the specified area and therefore the flux density will be given by the flux divided by the area.

Think About It

If the magnetic flux density at surface A_1 is $\mathbf{B_1}$ =0.4 T then what is the magnetic flux density at A_2?



Solution

Recall that if the ferromagnetic material forms a closed loop the flux will preferentially travel within the ferromagnetic material. Therefore, the flux is constant throughout the ferromagnetic material. Think of the flow of flux like water flowing through both a wide and a narrow tube. If the water is incompressible then the same amount of water must flow through the total cross sectional area in the wide tube as will flow through the cross sectional area in the narrow tube. Therefore:

$$\vec{B}_1 = \frac{\Phi}{A_1} = 0.4\,\text{T} \Rightarrow \Phi = \vec{B}_1 A_1 \qquad \text{(x.x)}$$

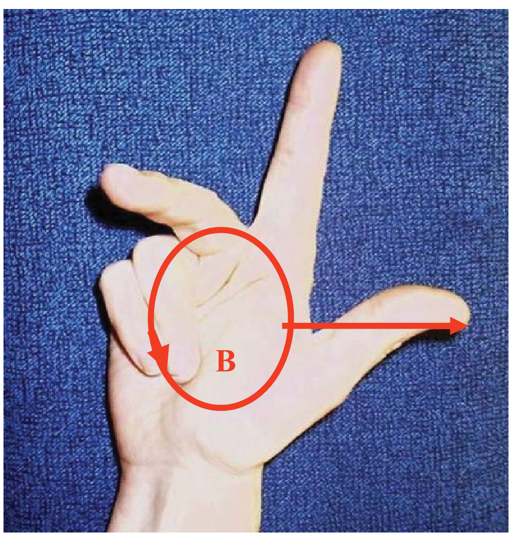
$$\vec{B}_2 = \frac{\Phi}{A_2} = \frac{\vec{B}_1 A_1}{A_2} \Rightarrow \vec{B}_2 = (0.4)\frac{0.02}{0.01} = 0.8\,\text{T} \qquad \text{(x.x)}$$

Practice Problems

1. Draw and label a bar magnet with the following: Height = 2 cm Width = 1 cm Length = 10 cm and the flux, Φ traveling S → N with a flux density, **B**= 0.14 T
2. What is the total flux in the magnet?

6.5 Electromagnet: Right Hand Rule

Remember, Øersted showed that a current passing through a conductor produces a magnetic field. Additionally, Faraday showed that the magnetic field is circular in nature. So how do we know the direction of the magnetic field? Taking your right hand, point the thumb in the direction of the current, the fingers will curl in the direction of the field lines (around the conductor)! Use Figure 6.5 as a guide to practice finding magnetic field direction with your own hand.



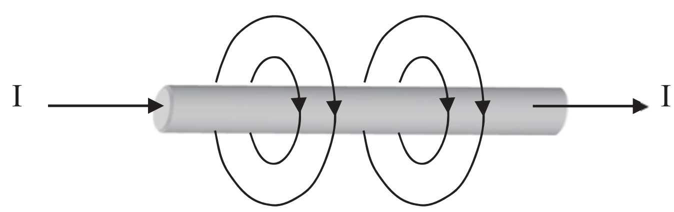
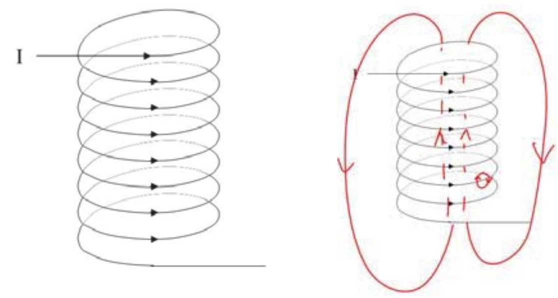


Figure 6.5: Illustration of the right hand rule for determining the direction of magnetic field.

To make an electromagnet we wrap an insulated wire around a ferromagnetic core. Each complete *wrap*, is a turn in the wire. All turns equal a *winding*. Using a cylindrical core, we can create a *solenoid* with a magnetic field and N and S poles. You can deduce the poles of the magnet from the direction of **B**. If we were to use a circular core we would create a *toroid* instead of a solenoid.

Think About It

Consider a coiled piece of wire like the one pictured here.
In what direction would the magnetic flux lines point?



Solution

To determine the direction of the flux lines for our coiled piece of wire we will use the right hand rule. Positioning our thumb in the direction of the current (I) ⇒, with our palm facing up, we can see that our fingers are curling in the direction of the field lines. Therefore the flux lines are traveling up, through the center of the coil of wire, drawn in on the following image.

6.6 Magnetic Circuits

We have seen so far how to find the direction of **B**. We will now consider the relationship between the current, I, in the windings and the produced flux in the core. A typical problem in magnetic circuits would be to determine the amount of current to be supplied to a winding to produce a specific flux in a core. If we take a ferromagnetic core and coil wire around it, the flux will be confined in the core creating an electromagnetic circuit. An example of this electromagnetic circuit is seen in Figure 6.6. (The shaded circle in Figure 6.6 represents the cross-sectional area (A) for the core.)

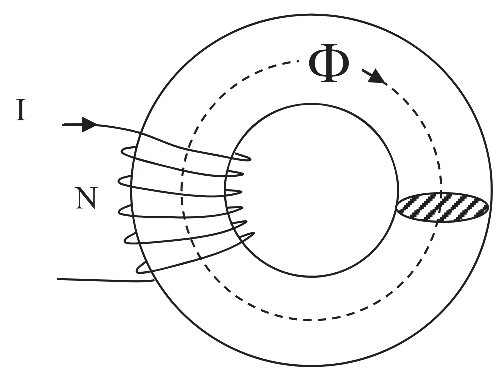


Figure 6.6: Illustration of a ferromagnetic core, μ, wrapped with wire; labeled with magnetic flux line (Φ), current (I), and the number of turns (N).

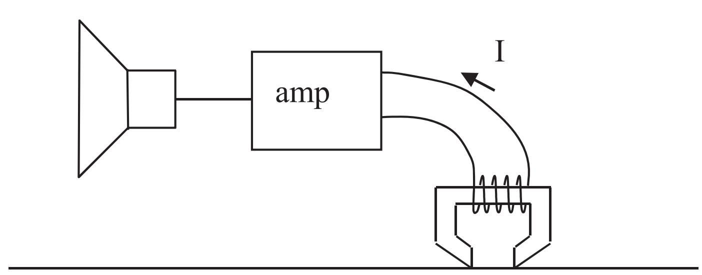


Figure 6.7: Illustration of a magnetic playback head.

By changing the current in the coil we can change the flux in the core. Remember that not only does current create magnetic flux, but changes in magnetic flux will also create a current. For example, Figure 6.7 illustrates how during playback of a tape recorder the magnetized pattern on the tape is detected by the playback head. The magnetic flux is confined to the head which then induces a current in the coil. This current is then amplified and sent to the speaker.

Consider a winding or coil around a core as seen in Figure 6.6. The winding is a current-carrying conductor. The flux producing ability of the coil is called its **magnetomotive force (mmf)**, $\Im$, which is defined as:

$$\Im = NI \qquad \text{(x.x)}$$

Each turn yields the same current and adds the same flux to the core. A coil with two turns will produce twice as much flux as a coil having only one turn. The magnetomotive force then is the amount of magnetic *excitation* applied to the electromagnet. More current produces more flux. A greater number of turns produces more flux.

Reluctance, $\Re$ is the opposition to magnetic flux that the circuit presents. In other words, the magnetomotive force is acting to produce flux, while the reluctance of the core is acting to oppose the flux. Therefore, we have:

$$\Im = \Phi\Re \qquad \text{(x.x)}$$

The reluctance of the core is defined in terms of the mean path length, L, in the core, the cross-sectional area, A, of the core, and the permeability, μ, of the core.

$$\Re = \frac{L}{\mu A} \qquad \text{(x.x)}$$

We can now see the analogy between electric and magnetic circuits.

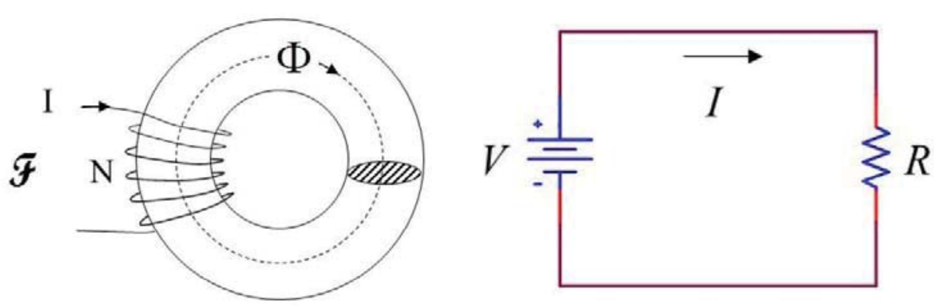


Figure 6.8: A comparison of magnetic circuit elements to those of electric circuit elements

Figure 6.8 shows the analogy between electric and magnetic circuits. The magnetomotive force ($\Im$), which is equal to the flux (Φ) times the reluctance ($\Re$), in the magnetic circuit is equivalent to Ohm's law, V = IR in the electric circuit. The mmf ($\Im$) takes the place of voltage (V), flux (Φ) takes the place of current (I), and reluctance ($\Re$) takes the place of resistance (R).

We can see the similarities in the definition of reluctance and resistance as well. Both are dependent on the mean length and inversely dependent on the cross sectional area. This makes for a good analogy, but it is not all that useful since reluctance depends on permeability, which depends on flux in ferromagnetic materials. This can make it difficult to solve for flux!

6.7 Magnetic Field Intensity

For simple geometries we can compute a magnetic field intensity produced by a current carrying conductor. The magnetic field intensity, **H**, is constant on a circle in which all points on it are equidistant from a wire carrying current, I.

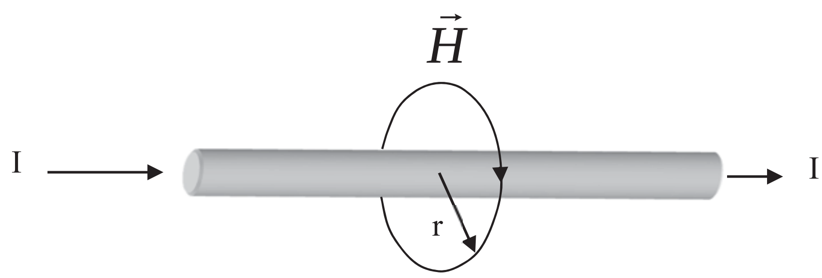


Figure 6.9: Magnetic field produced for a straight wire carrying a current.

Ampère discovered that the magnetic field intensity along a closed path is proportional to the current enclosed in the path and inversely proportional to the length of the path. In other words, the magnetic field intensity on a circular path, a distance 'r' from the current, is equal to:

$$\vec{H} = \frac{I}{2\pi r} \qquad \text{(x.x)}$$

where the 2πr is the circumference of the circle (the path around which the magnetic field exists).

If we now assume that the magnetic field is confined to the core of a toroid being driven by 'N' turns of a conductor carrying 'I' amps, as seen in Figure 7.17, then the total current enclosed is N times I and the mean path is the path in the core. Mathematically we have:

$$\vec{H} = \frac{NI}{2\pi r} = \frac{NI}{L} \qquad \text{(x.x)}$$

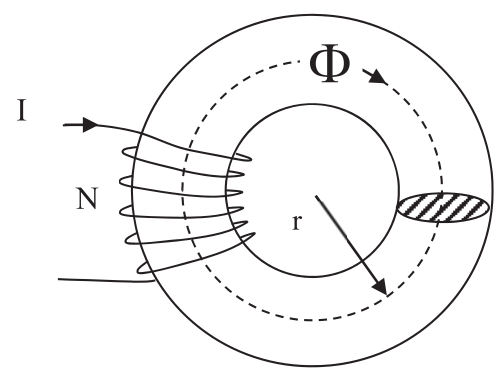


Figure 6.10: Geometry of a core showing magnetomotive force, NI, and mean path, L=2πr

Where L is the mean path in the core (of any geometry). From this equation we can see that the magnetic field intensity, **H**, is the measure of the magnetomotive force, NI, per unit length, L.

Ampère's Law*: The magnetic field intensity,* ***H*** *(A/m), along a closed path is proportional to the current enclosed in the path and inversely proportional to the length of the path.*

6.8 Ampère's Circuital Law

Rearranging equation (x.x) gives:

$$NI = \vec{H}L \qquad \text{(x.x)}$$

Written in this way shows that we can consider the NI as the magnetomotive *source*, while the HL is the magnetomotive *load* or *drop* (see Figure 6.11).

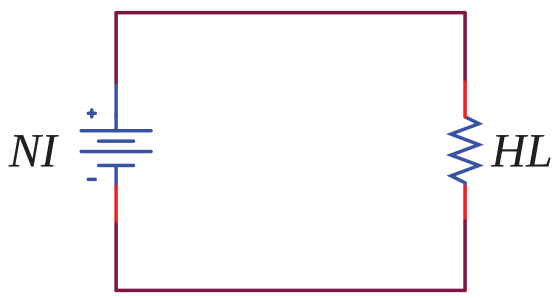


Figure 6.11: Model of magnetic circuit elements using electric circuit analogy.

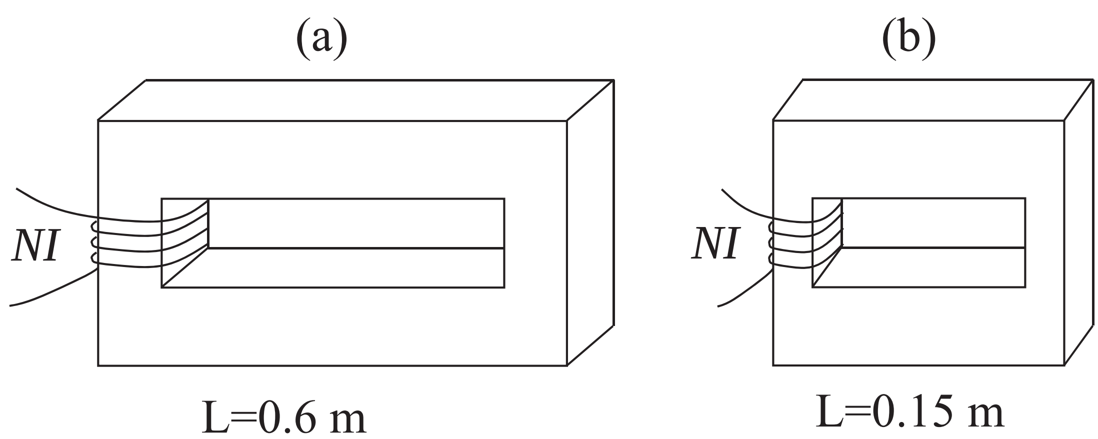
This looks like Kirchhoff's voltage law applied to magnetic circuits. In fact, if there are multiple H's along the path, for instance if there are two different materials in the path, then we have:

$$NI = \sum_i \vec{H}_i L_i \qquad \text{(x.x)}$$

***Ampère's Circuital Law** states that the sum of the magnetomotive forces around a magnetic circuit equals the sum of the magnetomotive drops around that circuit.*

Think About It:

Compare circuits (a) and (b):



If NI is the same for both circuit (a) and circuit (b), which has the larger magnetic field?

Solution:
In both circuits (a) and (b), we have:

$$NI = \vec{H}L$$

But in (a) L is larger. Therefore, **H** must be smaller in (a) than in (b). (Since NI is the same for both.)

Practice Problems

1. How much current must flow in a 15-turn winding to establish a magnetic field intensity of 60 A/m in a circular toroid (r=3 cm)?
2. How many turns would be required to establish such a magnetic field intensity in the toroid if the current supplied to the winding is 500 mA? (Answer to the closest integer.)

6.9 Permeability and B-H Curves

Recall that **H** is a measure of the magnetizing force (mmf per unit length) that produces flux (**Φ**). Thus:

$$\vec{H} = \frac{\Im}{L} = \frac{NI}{L} \qquad \text{(x.x)}$$

Now consider **B**, which is a measure of the resulting flux, **Φ**. Where **Φ** is divided by the cross sectional area (A):

$$\vec{B} = \frac{\Phi}{A}\hat{n} \qquad \text{(x.x)}$$

We can see that both **H** and **B** are related to flux. They must both then be related to each other. Rearranging the first two equations and then substituting for the definition of reluctance gives:

$$\Phi = \frac{\Im}{\Re} \Rightarrow \vec{B}A = \frac{\vec{H}L}{L/\mu A} \Rightarrow \vec{B} = \mu\vec{H} \qquad \text{(x.x)}$$

From this we see that the magnetic flux density, **B**, is related to the magnetic field intensity, **H**, by the permeability, μ, of the material that contains the flux.

For free space the ratio of magnetic flux to the magnetic field is a constant. The permeability of a vacuum (or free space) is:

$$\mu_o = \frac{\vec{B}_{air}}{\vec{H}_{air}} = 4\pi\times10^{-7}\,[\mathrm{Wb/A\cdot m}] \qquad \text{(x.x)}$$

Generally, however, permeability is not a constant. It depends on the amount of flux in the material.

Ferromagnetic materials have the ability to retain magnetism after they have been subject to a magnetic field. That is how magnetic tape works. This ability to remember the magnetic field, even when it is no longer present, is called *hysteresis*. Additionally, ferromagnetic materials will saturate at very large flux. The result is that the ratio of the magnetic flux to the magnetic field in a ferromagnetic material is not constant but depends on the amount of flux in the material.

$$\mu_{ferro} \neq \frac{\vec{B}_{ferro}}{\vec{H}_{ferro}} \qquad \text{(x.x)}$$

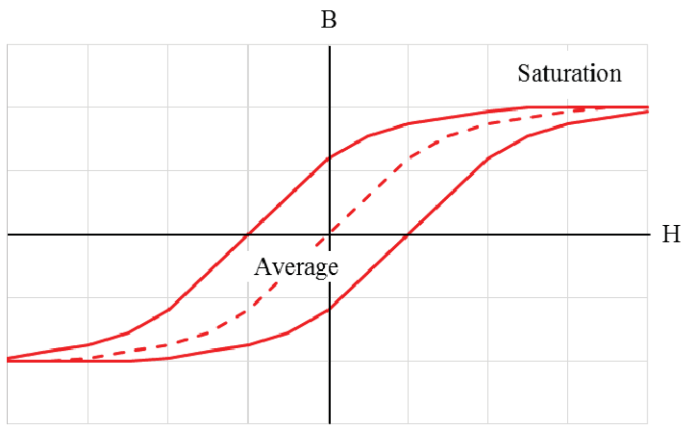


Figure 6.12: Example B-H curve showing hysteresis and saturation

Figure 6.12 shows an example curve that relate magnetic flux density, **B**, to the magnetic field, **H**. The hysteresis is represented by the double curve. Saturation is represented by the curve reaching a maximum magnetic flux density where it is no longer dependent on magnetic field. So rather than providing the permeability of a ferromagnetic material, most materials are defined in terms of their average B-H curve or magnetization curve.

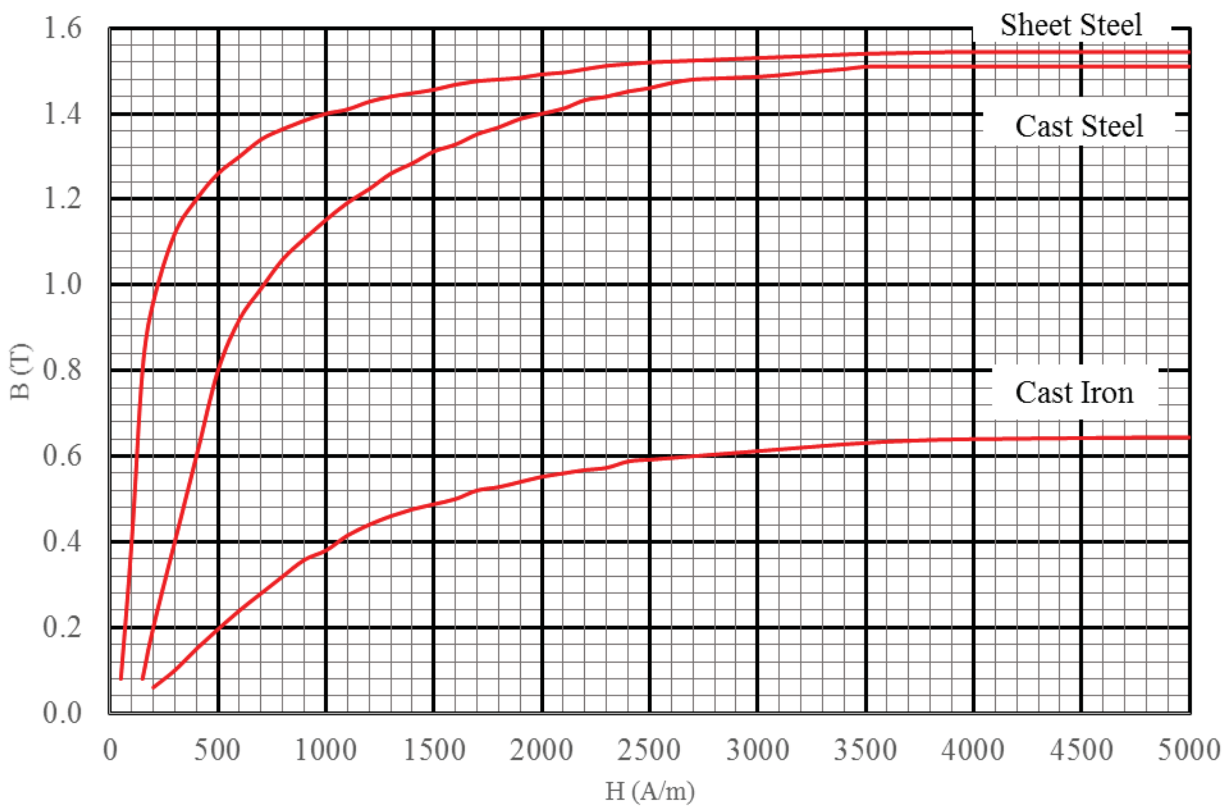
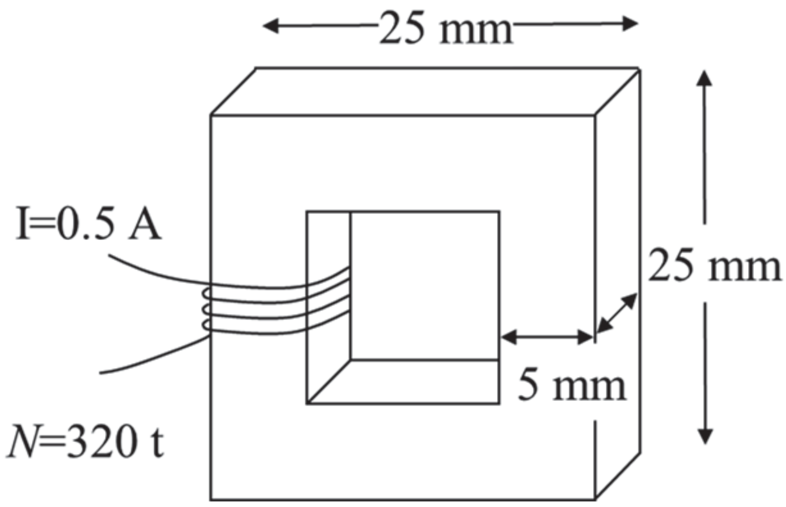


Figure 6.13: Average B-H Curves for three materials

The B-H curve for sheet steel, cast steel, and cast iron are shown in Figure 6.13. The B-H curve represents the "average" relationship between **B** and **H**. The hysteresis is not shown and we will effectively ignore it. We do, however, see a clear saturation at large values of **H**.

Think About It

What is the magnetic field intensity, **H**, magnetic flux density, **B**, and magnetic flux, Φ, in the following circuit? The material of the core is cast steel.



Solution

The magnetic field intensity, **H**, is found using the formula:

$$\vec{H} = \frac{NI}{L} \Rightarrow NI = \vec{H}L$$

We are given:
N = 320 t
I =0.5 A
L =4 ∗ (20 mm) (actual dimensions 25 by 25, mean path is 4*20)

Thus we can calculate our **H** directly as:

$$\vec{H} = \frac{NI}{L} = \frac{(320)(0.5)}{80 \times 10^{-3}} = 2000 \text{ A/m}$$

Using our magnetic field intensity value of 2000 At/m and the B-H curve graph shown in Figure 6.13, we can see that for the ferromagnetic material, *cast steel*, the B value is approximately 1.4 T.

B =1.4 T

Finally we need to solve for the magnetic flux, Φ, in the circuit. We can do this by using our previously solved for magnetic flux density, **B**, and the cross sectional area of our circuit. We also know that **B** is equal to the ratio of the flux over the cross sectional area.(**B** =Φ/A). Rearranging this equation we have:

$$\Phi = \vec{B}A = (1.4)(5\times10^{-3})^2 = 35\times10^{-6}\ \text{Wb}$$

Practice Problems

1. Find the current (I) that must be supplied to a 50-turn winding in a toroid in order to produce a flux of 3×10^{-5} Wb in the core. The cross-sectional area of the cast steel core is 4×10^{-5} m^2 throughout its length. The radius of the toroid is 3 cm. (Use μ c 1.596×10^{-3} H/m)

7.10 Electromagnet with an Air Gap

The rotation speed of motors increases with the flux density **B**. An air gap must exist to allow some parts of the motor structure to rotate. We will show in the following section why the air gap must be kept small as possible for the motor to be efficient (indeed **B** decreases with the size of the air gap).

If we have a magnetic circuit that contains more than one material, i.e. the material of the core and the material of the air, then we need to calculate the **H**L for each material. Remember that Ampère's Circuital Law states that the sum of the magnetomotive forces, NI, equals the sum of the magnetomotive drops, **H**L.

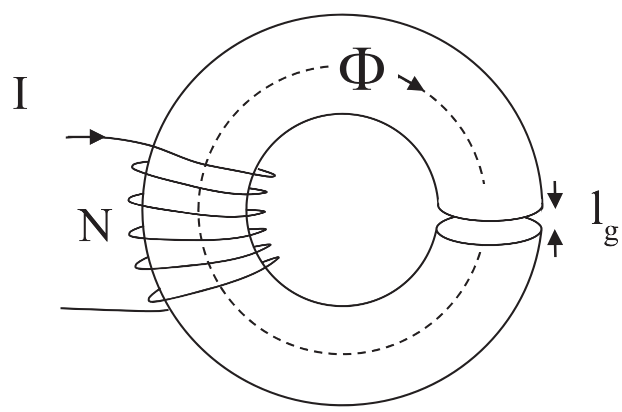


Figure 6.14: Electromagnet with an air gap.

There are two materials for the circuit shown in Figure 6.14, the core material and air. Applying Ampère's Circuital Law gives:

$$mmf_{\text{supplied}} = mmf_{\text{air}} + mmf_{\text{core}}$$

or

$$NI = l_g H_{air} + (l - l_g) H_{core} \tag{x.x}$$

where l_g is the length of the gap and l is the mean path around the entire magnetic circuit. Therefore l-l_g is the length of the path in the core. Relating the magnetic field to the magnetic flux gives:

$$NI = l_g \frac{\vec{B}_{air}}{\mu_{air}} + (l - l_g)\frac{\vec{B}_{core}}{\mu_{core}} \tag{x.x}$$

Remember that the magnetic flux density, **B**, is just the total flux, **Φ**, divided by the cross-sectional area, A. The magnetic circuits discussed in this class are all "series" circuits, therefore the flux is constant throughout the circuit (similar to the current I being constant in a series electrical circuit). When fringing effects, (see Figure 6.15) can be ignored the cross sectional area of the air gap

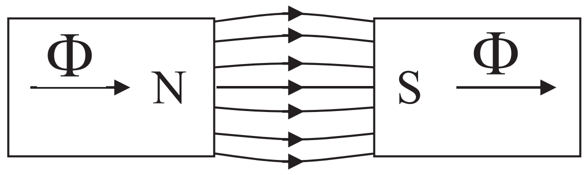


Figure 6.15: Illustration of an air gap with negligible fringing effects.

equals the cross sectional area of the core at the air gap. If both the flux and the cross-sectional area are constant then the flux density **B** must also be constant, i.e. $\vec{B}_{airgap} = \vec{B}_{core}$.

Finally, since the permeability of air is so small, most of the magnetomotive force drop is across the air gap. Mathematically we have:

$$NI = \left(\frac{l_g}{\mu_o} + \frac{(l - l_g)}{\mu_{core}} \right) \vec{B} \Rightarrow NI \approx \frac{l_g}{\mu_o} \vec{B} \qquad \text{(x.x)}$$

Recall that the permeability of a vacuum (or free space) is:

$$\mu_o = \frac{\vec{B}_{air}}{\vec{H}_{air}} = 4\pi \times 10^{-7} \left[\mathrm{Wb/A \cdot m} \right] \qquad \text{(x.x)}$$

Think About It

Consider Figure 6.14 where the core is made of sheet steel, the cross-section of the core is circular with a radius of 1.25 mm, the mean length around the circuit is l = 30 cm, and the airgap, l_g = 1 mm.

1. Find the magnetomotive force that must be applied by the winding in order to develop a flux of 0.6×10^{-5} Wb throughout the core. Neglect fringing in the air gap.
2. Find the number of turns in the winding, assuming a current of 40 A.
3. What would the number of turns be in the winding (to develop the same flux) assuming no air gap?

Solution

1. To find the magnetomotive force that must be applied by the winding in the figure in order to develop a flux of 0.6×10^{-5} Wb throughout the core we must first recognize that $\vec{B}_{airgap} = \vec{B}_{core}$. If the cross sectional area of the materials (core and air) are the same then the measure of the resulting flux, B, is a constant.

$$\vec{B} = \frac{\Phi}{A} = \frac{0.6 \times 10^{-5}}{\pi \left(1.25 \times 10^{-3}\right)^2} = 1.222$$

 Use Ampère's Circuital Law to solve for our magnetomotive force, NI.

$$NI = l_g H_{air} + (l - l_g) H_{core}$$

 In order to solve for NI, we first need to solve for the magnetic field intensity **H** in both the air and the core. The relationship between **B** and **H** in air is given by:

$$\vec{B}_{air} = \mu_o \vec{H}_{air}$$

where $\mu_o = 4\pi \times 10^{-7}$. Solving for $\mathbf{H_{air}}$ gives:

$$\vec{H}_{air} = \frac{\vec{B}_{air}}{\mu_o} = \frac{1.222}{4\pi \times 10^{-7}} = 972{,}436.7 \text{ At/m}$$

Note that the unit At/m (amp turns/ meter) emphasizes the fact that typically magnetic field will be generated by a coil of N turns carrying I amps. The magnetic field intensity in the core is determined reading the value off the B-H curve for sheet steel. Using B=1.222 and Figure 6.13, we find

$$\vec{H}_{core} = 450 \text{ At/m}$$

Plugging our values of $\mathbf{H_{core}}$ and $\mathbf{H_{air}}$ into Ampere's Circuital Law gives:

$$\begin{aligned} NI &= l_g H_{air} + (l - l_g) H_{core} \\ &= \left(1\times 10^{-3}\right)\left(972{,}436.7\right) + \left(30\times 10^{-2} - 1\times 10^{-3}\right)\left(450\right) \\ &= 1106.9867 \text{ At} \end{aligned}$$

2. The number of turns required for a coil carrying a current of 40 A is:

$$N = \frac{NI}{I} = \frac{1106.9867}{40} = 27.67 \text{ or } 28 \text{ turns}$$

3. Without the airgap we only have one material, so Ampere's Circuital Law reduces to:

$$NI = lH_{core} \Rightarrow N = \frac{lH_{core}}{I} = \frac{\left(30\times 10^{-2}\right)\left(450\right)}{40} = 3.375 \text{ or } 4 \text{ turns}$$

6.10 Electromagnetic Machines

Magnetic forces act on moving charges. The force equation states that the force is equal to the cross product between the charge, q, times its velocity vector, $\boldsymbol{v}$, and the magnetic flux density, $\boldsymbol{B}$, producing the force.

$$\vec{F} = q\vec{v} \times \vec{B} \qquad \text{(x.x)}$$

Pay attention to the fact that the × symbol in equation (x.x) indicates that cross-product, not multiplication. The implication of the cross product is that the force is perpendicular to both the velocity vector of the charge and the magnetic flux. If the charge was stationary or moving parallel to the magnetic field then the magnitude of the force would be zero.

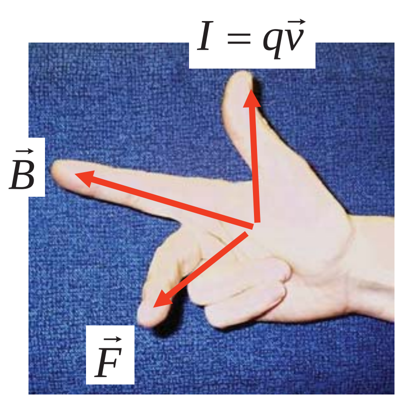


Figure 6.16: Right Hand rule for magnetic force

We can determine the direction of the magnetic force by using a variation of the right hand rule. Figure 6.16 illustrates how to use your right hand in determining vector directions. To do this, point your thumb in the direction of the charge velocity vector and your index finger in the direction of the magnetic flux. Your middle finger will then be pointing in the direction of the force. We note the measure of moving charge is a current.

If the magnetic force relationship is applied to a current-carrying wire then the force is the cross product between the direction of the current and the magnetic flux.

$$\vec{F} = I\vec{L} \times \vec{B}$$

Again we can apply the right hand rule to determine the force on the wire. Figure 6.16 illustrates how the graphical vectors are related to the right hand rule for reference.

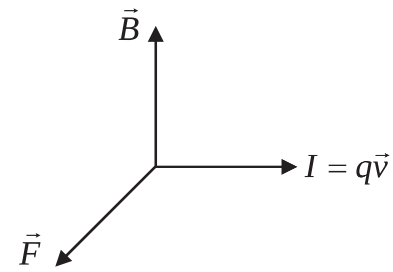


Figure 6.17: Vector equivalent to the right hand rule.

Consider an oscilloscope or other cathode ray tube, like the one illustrated in Figure 6.18. The cathode is held at a negative potential (-V) and ejects an electron into a magnetic field that points into the plane of the page ⊗. Will the electron travel up or down?

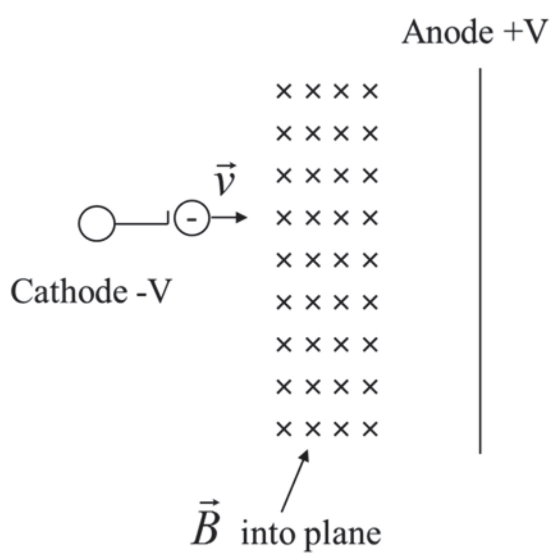


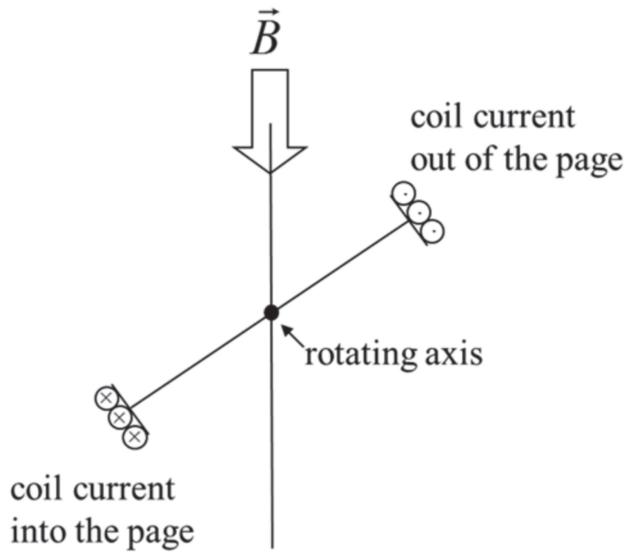
Figure 6.18: Illustration of a cathode ray tube showing the magnetic flux density going into the plane of the page

Using the right-hand rule illustrated in Figure 6.16 and Figure 6.17 we can determine that the force on the electron is pointing down. Be careful that equation x.x is for a positive charge. Newton's second law says that $\vec{F} = m\vec{a}$. In other words the electron will accelerate in the direction of the force. Therefore, the electron will be deflected downward.

Note: Describing magnetic force requires a discussion of three dimensions. In order to describe a three dimensional space on a two dimensional piece of paper we will use '⊗' to indicate into the page and ⊙ to indicate out of the page. Think about an arrow, →. The tip of the arrow, out of the page is the ⊙ and the back of the arrow is the ⊗.

Think About It

Consider the image below which shows a cross-section of a coil of wire that is going into and out of the plane of the page. A coil of wire contains a current and is permitted to rotate about an axis. The coil of wire is called the rotor. What happens to this coil when it is placed in a magnetic field as shown? To analyze this electromagnetic machine we need to ask:



1. What is the direction of the force felt by the current on the right side?
2. What is the direction of the force felt by the current on the left side?
3. What about the rotating force? Does the rotor rotate?
4. Does the rotor turn clockwise or counter clockwise?
5. Is there a stable position?

Solution

1. Consider the right hand side where the current is coming out of the page. Using the right hand rule, we point our thumb in the direction of the current coming out of the page, our index finger in the direction of the magnetic field, down, and finally our middle finger now points in the direction of the force, outward (→).
2. On the left hand side the current is going into the page. Again, the right hand rule says, thumb in the direction of the current, into the page, index finger in the direction of the magnetic field, downward, middle finger points in the direction of the force, again outward (←).
3. So the force on both sides of the rotor is pulling outward. Given its current position to rotor will rotate clockwise until it reaches the horizontal position.
4. The rotor will turn clockwise.
5. The forces on the rotor are always outward, so when the rotor passes the horizontal position the outward force will pull the rotor back to the horizontal position. The rotor will then oscillate around this position until it stabilizes.

If we take the rotor from the "Think About It" section and add a commutator, illustrated in Figure

6.19, we can reverse the current at suitable times, causing the previously stable position of the previous rotor to become unstable.

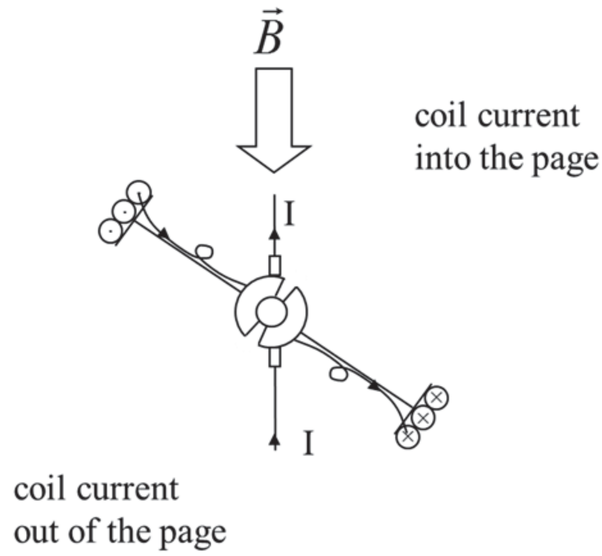


Figure 6.19: Rotor with a commutator

Initially the two rotors will act in the same fashion. The current out of the page will be forced right/down, and the current into the page will be forced left/up. However, when the rotor passes through the horizontal, the commutator reverses the current such that the coil on the left will now be out of the page and feel a force to the right/up while the coil segment of the right will now be into the page and feel a force to the left/down. Therefore, the rotor will continue to rotate in the clockwise direction.

The idea of using magnetic forces to create a rotation machine is the basis of motors and generators. A motor transforms electrical power into mechanical power while a generator transforms mechanical power into electrical power.

Chapter 7: I-V Characteristics

A plot of the current (I) versus voltage (V) of a component or sub-circuit gives you all the information you need to know about the electrical behavior of the component or circuit. In this chapter you will learn how to determine the I-V characteristics of basic components (e.g. resistors, voltage and current sources) and sub-circuits. You will also learn how to determine the Thévenin equivalent circuit for any given I-V characteristic. You will learn how to use a sub-circuit's I-V characteristics to determine the operating point when connected to a second sub-circuit.

7.1 Key Concepts

- Current and voltage are <u>dependent</u> quantities.
- Ohm's law specifies the <u>dependence</u> of voltage and current for a resistor, i.e. V/I is a <u>constant</u> ratio equal to the resistance R.
- Other devices have different <u>dependencies</u> between the current and voltage.
- Knowing the <u>dependence</u> of I versus V for any device or sub-circuit, S, gives you **ALL** the information you need to know about the behavior of that device or sub-circuit.

7.2 I-V Characteristics of Components

Consider a sub-circuit that consists of a single 500 Ω resistor shown in Figure 7.1(a). We can determine the I-V characteristics for this resistor by connecting an arbitrary source at the input and using Ohm's Law, $I = V/R$. Plotting *I* as a function of *V* gives Figure 7.1(b).

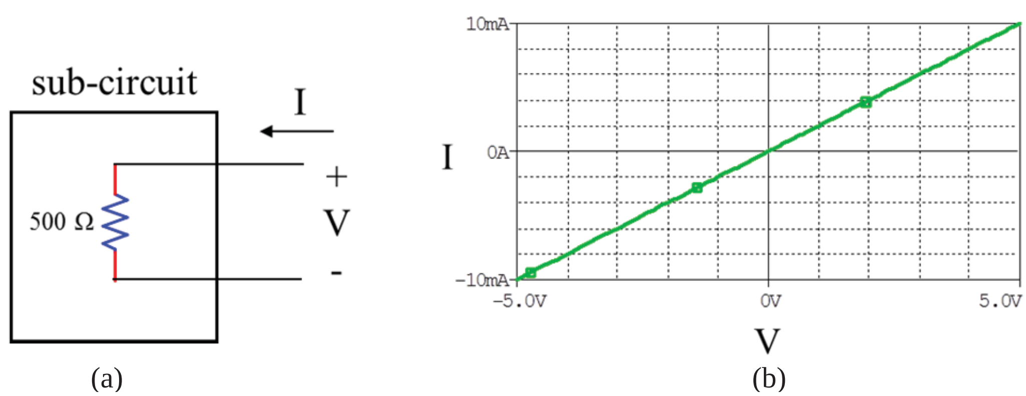


Figure 7.1: Example (a) resistor sub-circuit and (b) its I-V charactersitics.

Note that Figure 7.1(b) is a straight line, that the y-intercept is zero and the slope is 1/R. In this section we will consider only linear sub-circuits, i.e. sub-circuits for which the I-V characteristics are linear equations. In other words, $I = mV + b$, where m is the slope of the line and b is the y-intercept.

Two sub-circuits are *equivalent* if they have the same I-V characteristics. For instance, if I couldn't see inside the box of Figure 7.1(a) and could only measure the graph shown in Figure 7.1(b) then I wouldn't really know if what was inside the box was a 500 Ω resistor, two 1 kΩ resistors in parallel or two 250 Ω resistors in series. Each circuit would give the same I-V characteristic and are therefore *equivalent*.

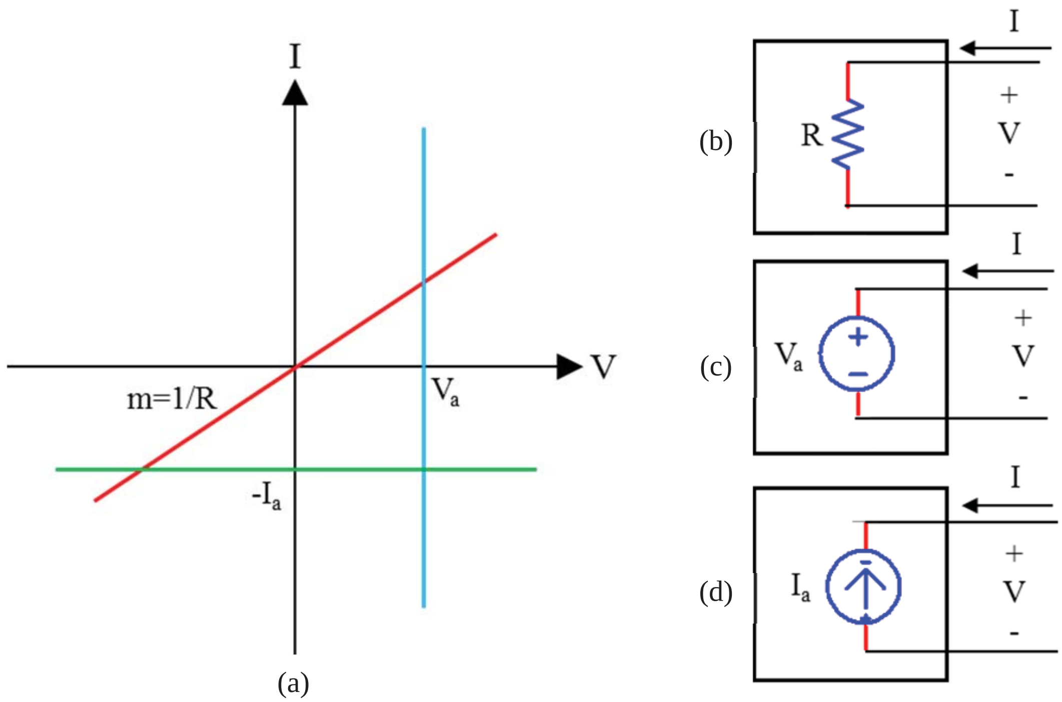


Figure 7.2: I-V characteristics for three components, (b) resistor – red line, (c) voltage source – blue line, and (d) current source – green line

Consider the I-V characteristics shown in Figure 7.2(a). The three lines on the graph represent the three different components shown in Figure 7.2(b-d). As we have previously seen a resistor has a linear I-V curve that crosses the origin and has a slope equal to 1/R. The red line represents the I-V curve of the resistor. An ideal voltage source is one that maintains a constant terminal voltage no matter how much current is drawn from it. The vertical blue line represents the voltage source. An ideal current source will supply the same current to any load connected across its terminals. The horizontal green line represents the current source. Remember that the measurement of the current and voltage for an I-V curve must be in passive sign convention with each other regardless of what is in the "box". So in this particular example the voltage source shown in Figure 7.2(c), has its I-V curve measured such that the voltage is in the same polarity as the actual supply, therefor the graph shows a vertical line crossing at $+V_a$. For the current source shown in Figure 7.2(d), the I-V curve is measured such that the current is in the opposite direction from the actual supply, therefore the graph shows a horizontal line crossing at $-I_a$.

I-V graphs are also useful for identifying when a component or sub-circuit is supplying power or dissipating power. If the I-V curve falls in the upper right hand quadrant or the lower left hand quadrant then the power is positive and the component or sub-circuit is dissipating power (see Figure 7.3). If the I-V curve falls in the upper left hand quadrant or the lower right hand quadrant then the power is negative and the component or sub-circuit is supplying power. Of course this interpretation is only valid if the I-V curve is measured in passive sign convention.

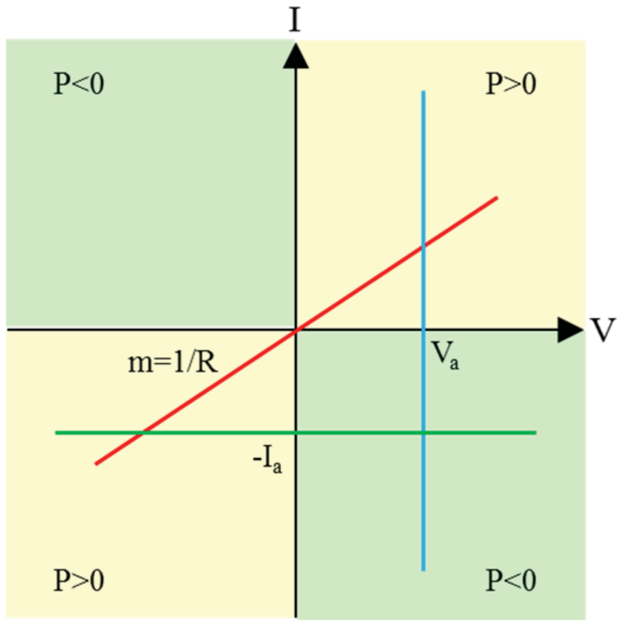


Figure 7.3: Illustration showing where power is dissipated and where power is supplied

Examining Figure 7.3, we see that the I-V curve for the resistor falls in the lower left hand quadrant and the upper right hand quadrant. Resistors always dissipate power. On the other hand the voltage source sometimes supplies power and sometimes dissipates power depending on what is connected to it. The same is true of the current source.

7.3 I-V Characteristics of Sub-Circuits

Consider Figure 7.5 which shows an I-V curve for an arbitrary sub-circuit. We can easily write the equation that represents this I-V curve by noting where it crosses the I and V axis. The point where the curve crosses the V axis (I=0) is called the open circuit voltage, Voc. The point where the curve crosses the I axis (V=0) is called the short circuit current, Isc (see Figure 7.4).

$$I = \frac{-I_{sc}}{V_{oc}} V + I_{sc} = \frac{-I_{sc}}{V_{oc}} \left(V - V_{oc} \right) \tag{7.1}$$

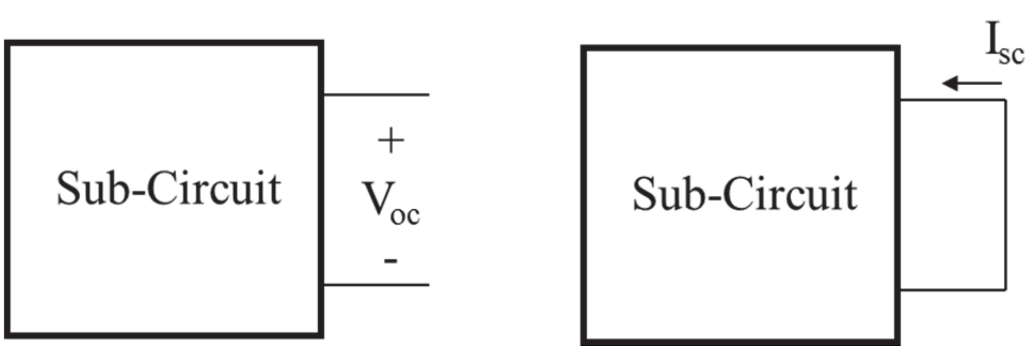


Figure 7.4: Measuring the open-circuit voltage and short circuit current.

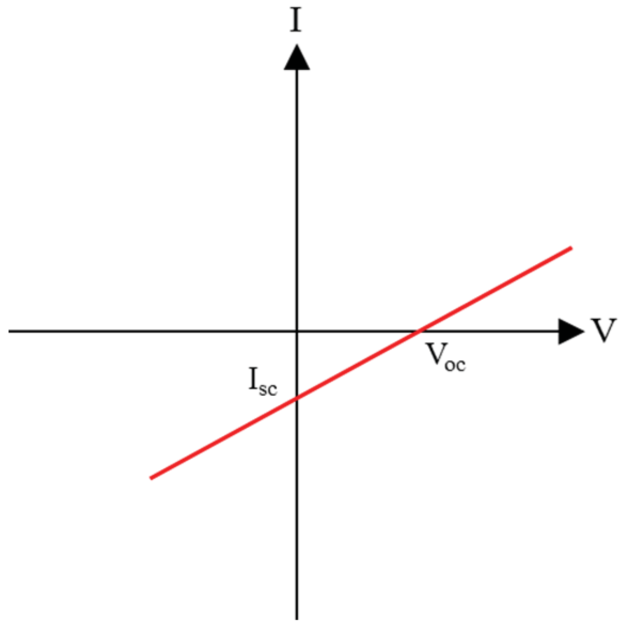


Figure 7.5: Example I-V curve of an arbitrary sub-circuit

From this we can derive a method to solve for the I-V curve of any arbitrary sub-circuit:

- Determine the open circuit voltage of the sub-circuit.
- Determine the short circuit current of the sub-circuit.
- Write the I-V equation (see 7.1).
- Draw the I-V curve.

Let's apply this method to solve the I-V characteristics of the circuit shown in Figure 7.6.

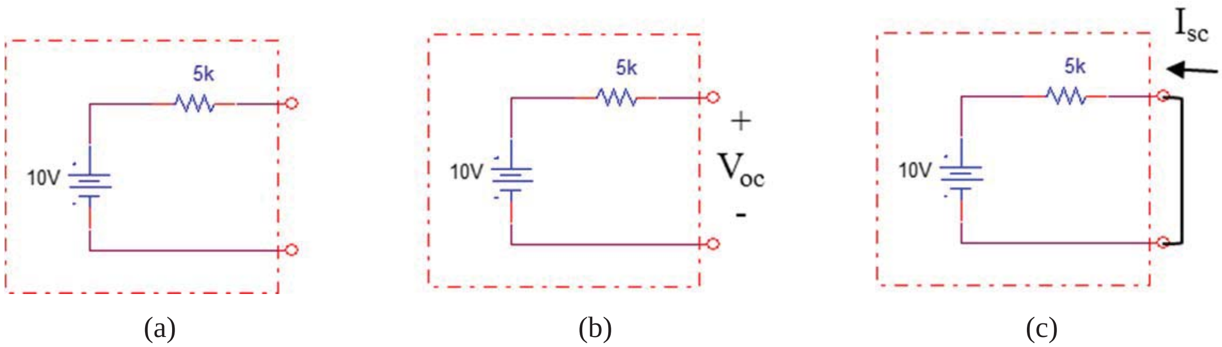


Figure 7.6: Simple circuit (a) showing open circuit voltage measurement (b) and short circuit current measurement (c).

Looking at Figure 7.6(b) we can see that the open circuit voltage is 10 V. The current through the resistor is zero and therefore from Ohm's law the voltage across the resistor is zero. Using KVL we can see that the voltage at the output terminals, V_{oc}, must equal 10 V. Looking at Figure 7.6(c) we can see that the short circuit current is -2 mA. Remember that the current through the resistor can be determined using the node voltages on either side of the resistor, i.e.

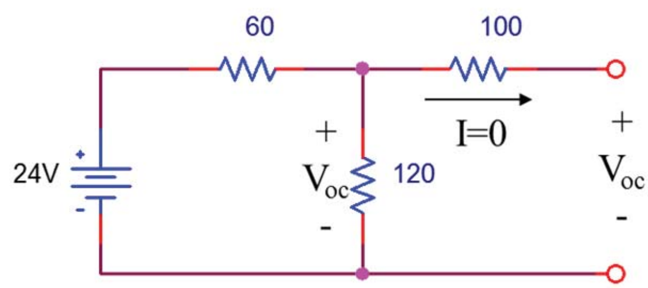
$$I_{sc} = \frac{0-10}{5000} = -2 \text{ mA} \tag{7.2}$$

$$I = \frac{2}{10}V - 2 = (0.2V - 2)\text{mA} \tag{7.3}$$

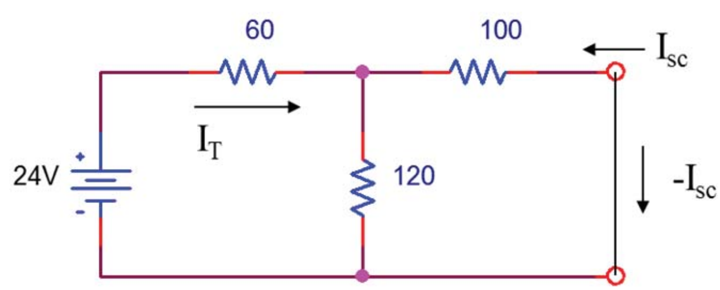
or

$$I = \frac{2}{10}(V-10) = 0.2(V-10)\text{mA} \tag{7.4}$$

A more complicated example is shown in Figure 7.7.



(a)



(b)

Figure 7.7: Second example circuit showing open circuit voltage measurement (a) and short circuit current measurement (b).

Using the same technique as before, we first solve for the open circuit voltage using Figure 7.7(a).

The current through the 100 Ω resistor equals zero and therefore there is no voltage drop across the 100 Ω resistor. Using KVL we can see that the voltage across the 120 Ω resistor is equal to Voc. We can solve for the voltage across the 120 Ω resistor by first determining the current through the resistor:

$$V_{oc} = I_{120}(120) = \left(\frac{24}{60+120}\right)(120) = 16\ \text{V} \tag{7.5}$$

Next we solve for the short circuit current using Figure 7.7(b). The easiest way to determine Isc is to first determine the equivalent resistance seen by the 24 V source and use that to determine I_T.

$$R_{eq} = 60 + \frac{1}{\frac{1}{120}+\frac{1}{100}} = 114.545\,\Omega \tag{7.6}$$

Therefore,

$$I_T = \frac{V_T}{R_{eq}} = \frac{24}{114.545} = 0.2095\ \text{A} \tag{7.7}$$

We can solve for the current through the 100 Ω resistor by first determining the voltage across the resistor:

$$I_{100} = \frac{V_{100}}{100} = \frac{I_T\left(\frac{100\cdot 120}{100+120}\right)}{100} = I_T\left(\frac{120}{100+120}\right) = (0.2095)\left(\frac{120}{100+120}\right) = 0.1143\ \text{A} \tag{7.8}$$

At this point you must be careful and ask yourself in what direction is the measured I_{100}? The direction of I_T implies that the polarity of the V_{100} measurement is positive on the top and negative on the bottom. This implies that I_{100} is measured in the direction of $-I_{sc}$. Therefore $I_{sc} = -0.1143\ \text{A}$. Finally, putting everything together we determine that the I-V characteristics for the circuit shown in Figure 7.7 is given by:

$$I = \frac{-I_{sc}}{V_{oc}}(V - V_{oc}) = \frac{0.1143}{16}(V-16) = 7.14(V-16)\,\text{mA} \tag{7.9}$$

Determining the open circuit voltage and the short circuit current is a straightforward way to determine the I-V characteristics of any arbitrary circuit. Unfortunately if you tried to actually measure the open circuit voltage and short circuit current on the bench you might damage the sub-circuit you are trying to measure. An alternative approach is to use nodal analysis. In the node method we:

- Attach a variable voltage source on the output of the sub-circuit (see Figure 7.8).
- Solve for I as a function of V using Nodal Analysis.

Consider the circuit shown in Figure 7.8. This is the same circuit we solved using the V_{oc}, I_{sc} method (see Figure 7.7). This time we will solve the circuit using Nodal Analysis. First we write KCL at node A.

$$I_{60} + I_{100} = I_{120} \tag{7.10}$$

$$\frac{24 - V_A}{60} + I = \frac{V_A}{120} \tag{7.11}$$

In writing equation 7.11, we recognize that $I_{100} = I$. Since we want the final equation to be of the form I equals some function of V, we don't substitute for I. Now I need some way to determine V_A in terms of V and/or I. Writing KVL around the right hand loop gives:

$$\begin{aligned} -V + I(100) + V_A &= 0 \\ V_A &= V - 100I \end{aligned} \tag{7.12}$$

Substituting back into equation (7.11) gives:

$$\frac{24 - (V - 100I)}{60} + I = \frac{(V - 100I)}{120} \tag{7.13}$$

Rearranging gives;

$$\begin{aligned} I &= \frac{(V-100I)}{120} - \frac{24-(V-100I)}{60} \\ I &= \frac{(V-100I)}{120} - \frac{48-2(V-100I)}{120} \\ 120I &= (V-100I) - 48 + 2(V-100I) \\ 420I &= 3V - 48 \\ I &= \frac{3}{420}V - \frac{48}{420} = \frac{3}{420}(V-16) \\ I &= 7.14(V-16) \end{aligned} \tag{7.14}$$

This is the exact I-V equation we derived in 7.9.

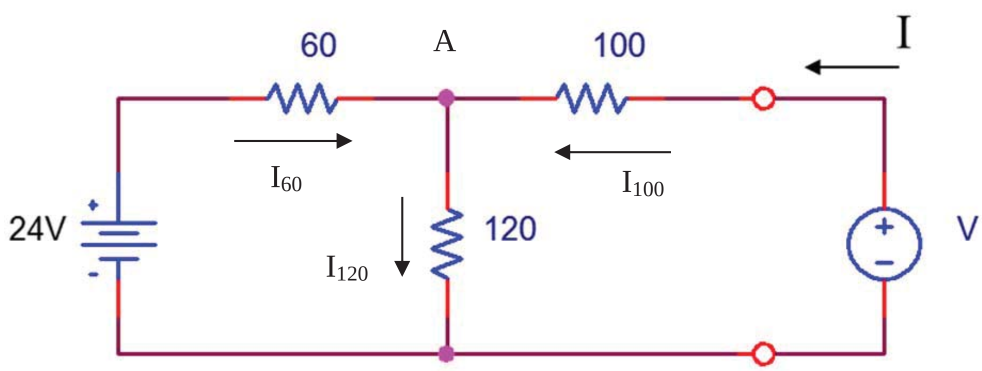


Figure 7.8: Solving for the I-V characteristics using Nodal Analysis

Once we have the I-V equation for any sub-circuit we can easily draw the I-V curve.

7.4 Thévenin Equivalence

Remember that if two circuits have the same I-V characteristics, they are equivalent. Figure 7.6(a) shows a very simple circuit which gave us the I-V equation:

$$I = 0.2(V - 10)\,\text{mA} \tag{7.15}$$

The more complicated circuit shown in Figure 7.7(a) gave a similar equation:

$$I = 7.14(V - 16)\,\mathrm{mA} \tag{7.16}$$

I would like to determine an equivalent circuit to Figure 7.7 that would give me the same I-V characteristics but that would look more like the simpler circuit shown in Figure 7.6. This equivalent circuit which consists of a DC voltage source in series with a resistor is called the Thévenin equivalent sub-circuit.

Remember that given any straight line with y-intercept equal to Isc and x-intercept equal to Voc, you can write the I-V equation as:

$$I = \frac{-I_{sc}}{V_{oc}}(V - V_{oc}) \tag{7.17}$$

As we have already determined, this straight line can be modeled with a single DC voltage source in series with a single resistor shown in Figure 7.9. By comparing this circuit with our solution to the circuit shown in Figure 7.6, we see that $V_{TH} = V_{oc}$ and $-I_{sc}/V_{oc} = 1/R_{TH}$. In other words the I-V characteristic equation for the circuit shown in Figure 7.9 is:

$$I = \frac{1}{R_{TH}}(V - V_{TH}) \tag{7.18}$$

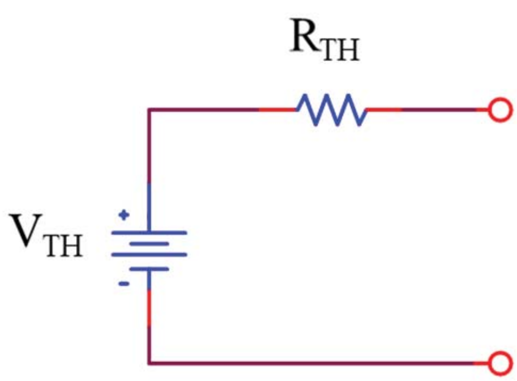


Figure 7.9: Thévenin equivalent circuit.

Going back to equation 7.16, we can immediately see that the Thévenin equivalent circuit for the circuit shown in Figure 7.7 is a 16 V source in series with a 140 Ω resistor as shown in Figure 7.10.

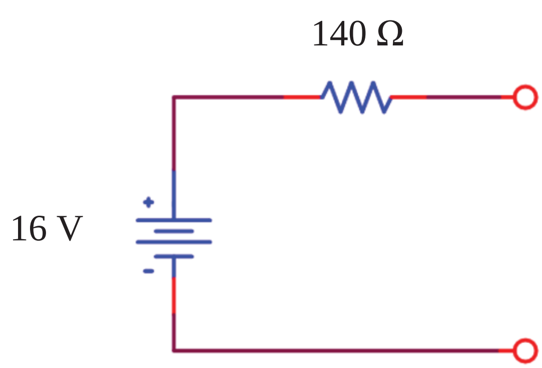


Figure 7.10: Thévenin equivalent circuit to Figure 7.7

You should convince yourself that the I-V characteristics for the circuit given in Figure 7.10 is the same as that in equation 7.16.

7.5 Circuit Analysis using I-V Curves

Consider two sub-circuits, A and B, connected together as in Figure 7.12. We are interested in determining the *operating point* when these two circuits are connected measured in the given polarity and direction of V_o and I_o. The only thing we know about these sub-circuits are their I-V curves shown in Figure 7.11.

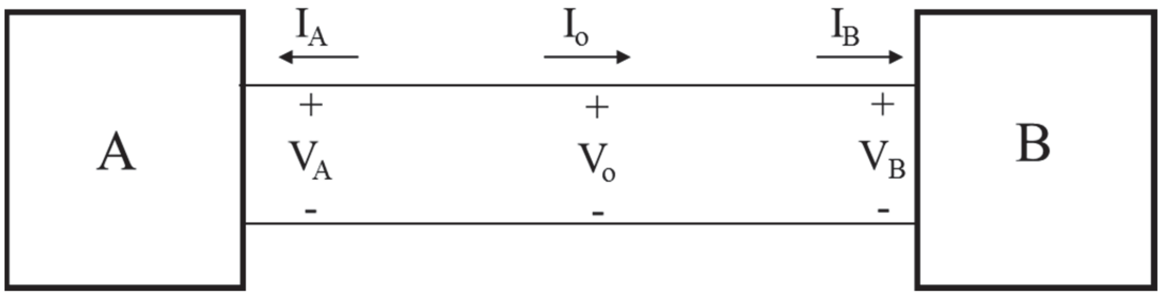


Figure 7.12: Circuit analysis of two connected sub-circuits

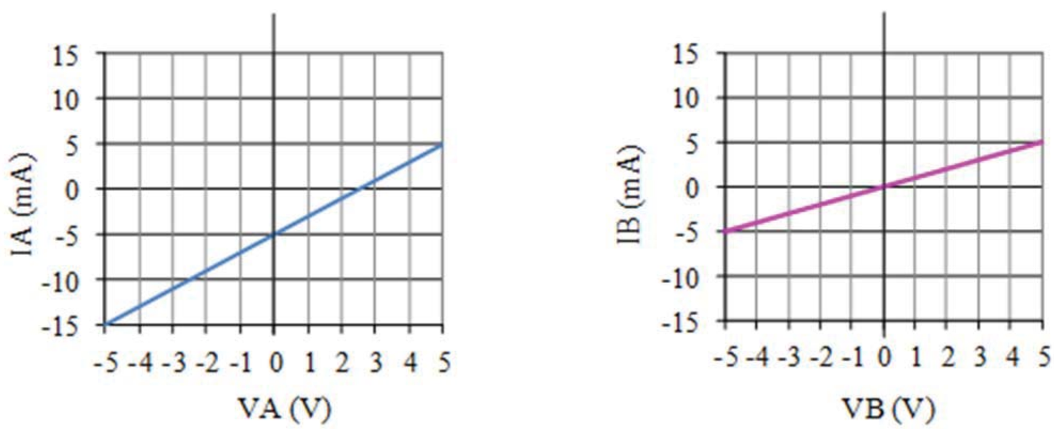


Figure 7.11: I-V curves for sub-circuits shown in Figure 7.12.

One method of solving for the operating point is to redraw the two graphs on top of each other after having translated them into the polarity and direction of V_o and I_o. For example, I_B=I_o and V_B=V_o, so drawing the I_B-V_B curve on an I_o-V_o graph will yield the same curve as shown in Figure 7.11 (see Figure 7.13 pink line). For sub-circuit A, $I_A = -I_o$ and V_A=V_o, so drawing the I_A-V_A curve on an I_o-V_o graph requires that the I_A-V_A curve be *flipped* about the V_A axis. Graphical *flipping* is equivalent to multiplying I_A by –1 (see Figure 7.13 blue line). The operating point for the two connected sub-circuits is the intersection of the two lines drawn in Figure 7.13, I_o≈1.6 mA, V_o≈1.6 V.

The graphical method provides an approximate solution to the operating point and is a useful check that a design is going to work as expected. A more accurate solution is found numerically by solving the simultaneous equations given by the two sub-circuit I-V equations. For our previous example, we can determine the I-V equations for sub-circuits A and B from their graphs in Figure 7.11.

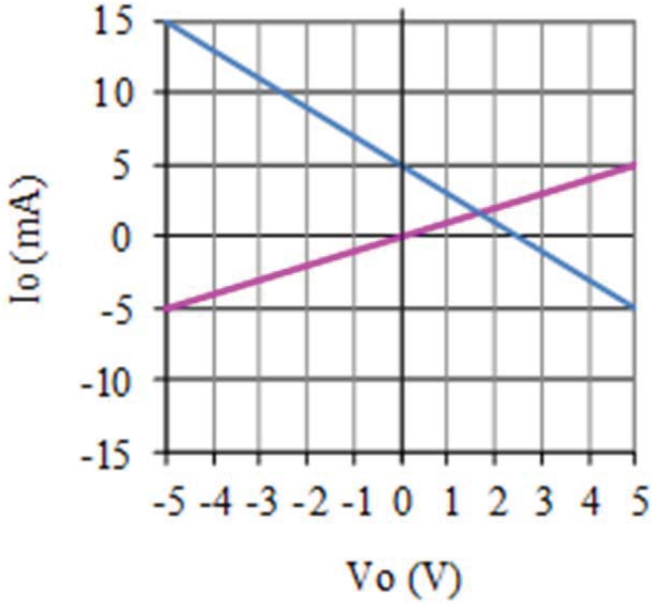


Figure 7.13: Graphical determination of operating point of two connected sub-circuits.

Solving for the I-V equations for sub-circuits A and B gives:

$$I_A = 2(V_A - 2.5)[\mathrm{mA}] \tag{7.19}$$

$$I_B = V_B\ [\mathrm{mA}] \tag{7.20}$$

I note that sub-circuit B must be a resistor circuit with an equivalent resistance equal to 1 kΩ. Sub-circuit A Thévenin equivalent circuit must be a 2.5 V voltage source in series with a 500 Ω resistor. You should convince yourself that this is true. In order to solve these two equations simultaneously, we need to recast them into equations of I_o and V_o (similar to our redrawing of the curves on the I_o-V_o graph). Remember, $V_o = V_A = V_B$, $I_o = -I_A = I_B$.

Equation 7.19 becomes:

$$-I_o = 2(V_o - 2.5)[\mathrm{mA}] \tag{7.21}$$

and equation 7.20 becomes:

$$I_o = V_o\ [\mathrm{mA}] \tag{7.22}$$

Now the two equations 7.21 and 7.22 can be solved simultaneously for I_o and V_o as follows:

$$
\begin{aligned}
-I_o &= 2V_o - 5 \\
I_o &= V_o \\
0 &= 3V_o - 5 \qquad (7.23) \\
V_0 &= \frac{5}{3} = 1.667\ \text{V} \\
I_o &= \frac{5}{3} = 1.667\ \text{mA}
\end{aligned}
$$

Chapter 8: Diodes

In this chapter we start our exploration of non-linear devices and how we can understand the operation of these devices through their piece-wise linear models. Specifically this chapter focuses on diodes, how they are created, how they operate (i.e. their I-V characteristics) and how we can analyze the operation of circuits that contain diodes. We start this chapter by discussing the properties of semiconductors and how those properties are changed when small amounts of impurities are added. This leads us to the discussion of the properties of a P-N junction and the diode I-V characteristics. Finally we develop a piece-wise linear model for the diode and use that model in the analysis of diode circuits.

8.1 Key Concepts

- **Diodes** are 2-terminal semiconductor devices with non-linear I-V characteristics.
- **Semiconductors** are crystalline solid materials with electrical properties midway between those of insulators and conductors.
- **Doping**, by adding impurities to a semiconductor, forms *n-type* or *p-type* materials which change the conductivity of the semiconductor.
- **P-N junction** is formed when a p-type material is connected to an n-type material. This forms a diode.
- **Biasing** establishes the proper operating conditions in electric components. The p-n junction operates such that current will flow in only one direction (*forward biased*) and not the other (*reverse biased*). The exception to this is Zener diodes which conduct current when reverse biased in the *breakdown* region.

8.2 Doping of Semiconductors

Remember that the electrons in the outermost shell of an atom are called valence electrons. The valence electrons dictate the nature of the chemical reactions of the atom and largely determine the electrical nature of solid matter.

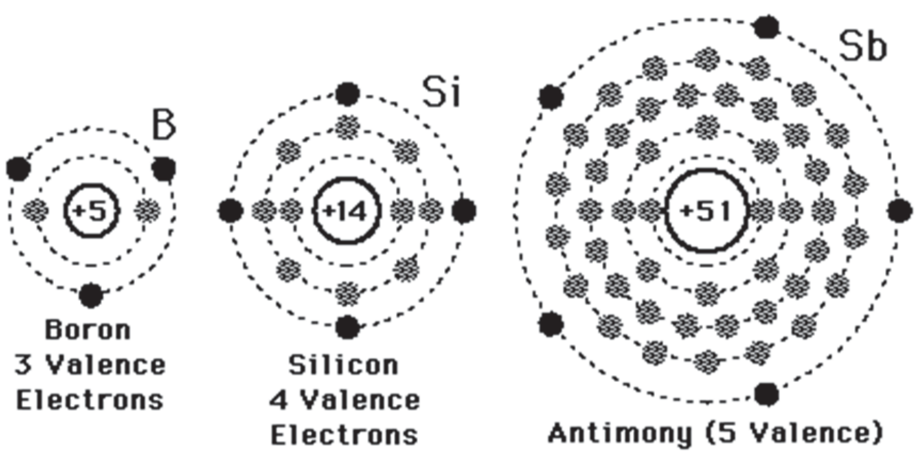


Figure 8.1: Boron, Silicon, and Antimony's atomic composition, most importantly their valence electrons, determine the electrial nature of the elements (http://hyperphysics.phy-astr.gsu.edu)

Figure 8.1 shows three different atoms, Boron, Silicon, and Antimony. Silicon is a common semiconductor used in everything from diodes and transistors to integrated circuits. It has four valence electrons. Boron and Antimony are common doping materials. They have one more or one less valence electron than silicon. Silicon atoms form covalent bonds and will crystallize into a regular lattice (see Figure 8.3). This crystal is called an *intrinsic* semiconductor and can conduct only a small amount of current. Intrinsic silicon acts like an insulator.

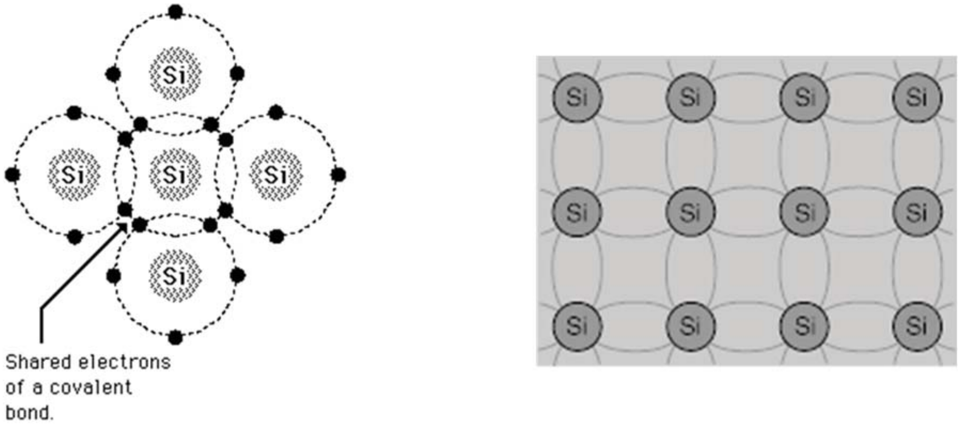


Figure 8.3: Silicon crystalline structure showing the shared electrons of a covalent bond. (http://hyperphysics.phy-astr.gsu.edu)

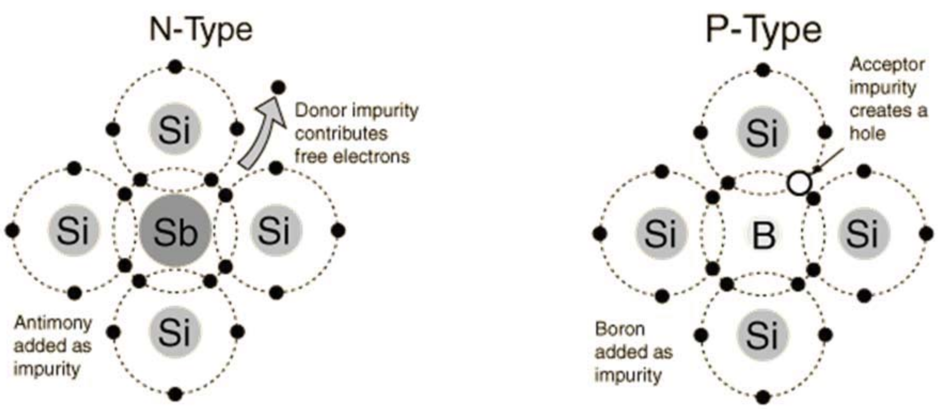


Figure 8.2: Example of how doping creates n-type and p-type semiconductors (http://hyperphysics.phy-astr.gsu.edu)

The addition of a small percentage of foreign atoms in the regular crystal lattice of silicon produces dramatic changes in their electrical properties. Figure 8.2 shows how silicon doped with Antinomy or Boron will create n-type or p-type semiconductors. For instance if the silicon is doped with Antinomy, which was five valence electrons, four of those electrons will form covalent bonds with the silicon crystal and one will be left free. This creates an n-type material, one with free electrons that are not bound to the crystal structure. Alternatively, if silicon is doped with Boron which has three valence electrons, then the silicon crystal will be left with a covalent bond that is unfilled. This unfilled location is called a *hole*. It is a place where if there was an electron nearby, the electron would try to fill the hole. Silicon doped with Boron creates a p-type material.

8.3 Properties of P-N Junctions

A p-n junction is formed by bringing together a p-type and an n-type semiconductor (see Figure 8.4). When a p-n junction is formed, some of the free electrons in the n-region diffuse across the junction and combine with the holes to form negative ions and leaving behind positive ions at the donor impurity sites. Remember that the n-type material is initially charge neutral. When the electrons diffuse across the boundary, the n-type material loses negative charge and therefore becomes positively charged. Likewise the p-type material is initially charge neutral. When the electrons diffuse across the boundary, the p-type material gains negative charge. After a while the space charge builds up, creating a depletion region which inhibits any further electron transfer unless it is helped by placing a forward bias on the junction.

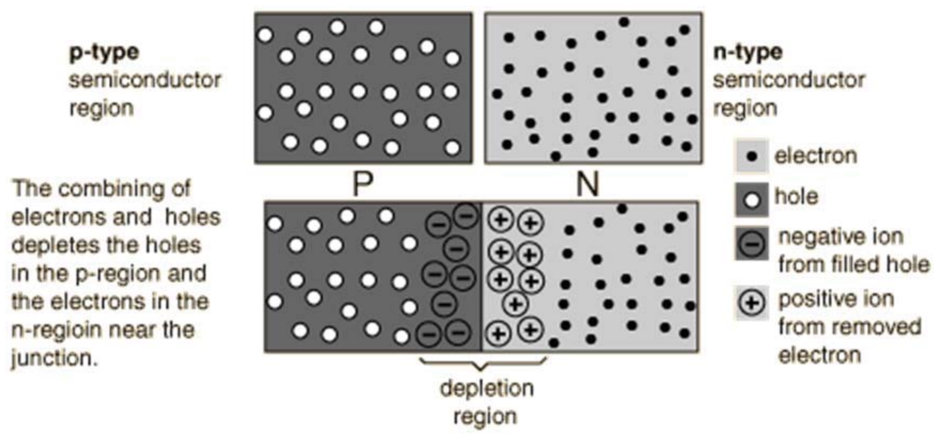
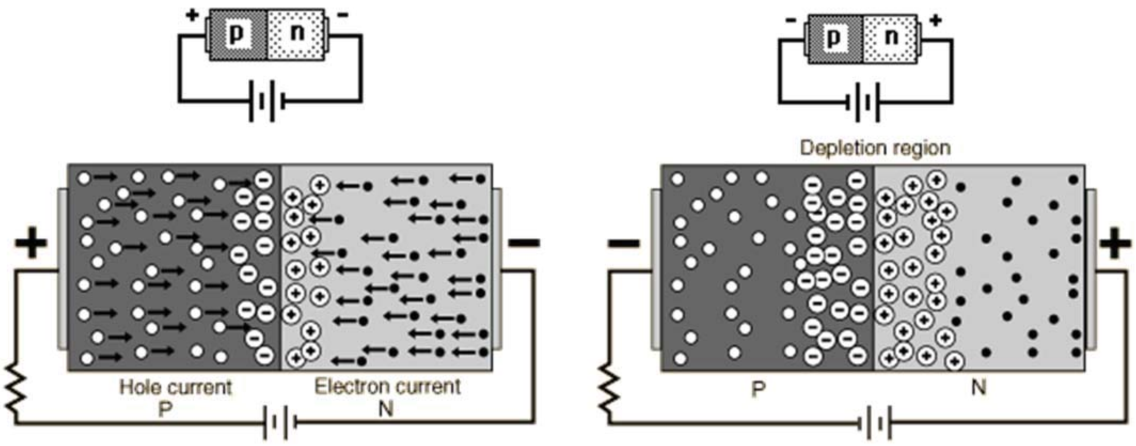


Figure 8.4: P-N junction showing depletion region

Forward biasing the p-n junction (see Figure 8.5(a)), by causing the p-type material to be at a more positive electric potential than the n-type material, drives holes into the junction from the p-type material and electrons into the junction from the n-type material. At the junction the electrons and holes combine so that a continuous current can be maintained. Reverse biasing the p-n junction (see Figure 8.5(b)), by causing the p-type material to be at a more negative electric potential than the n-type material, will cause a transient current to flow as both electrons and holes are pulled away from the junction. When the potential formed by the widened depletion layer equals the applied voltage, then the current will cease except for a very small thermal current.



(a) (b)

Figure 8.5: (a) Forward biasing of the p-n junction and (b) reverse biasing of the p-n junction. (http://hyperphysics.phy-astr.gsu.edu)

8.4 Diode I-V Characteristics

The nature of the p-n junction is that it will conduct current in the forward direction but not in the reverse direction. As seen by the I-V curve shown Figure 8.6, diodes have non-linear I-V characteristics. Take note of the current direction for I_D and the voltage polarity for V_D. Diode I-V curves are always defined with the voltage polarity and current direction.

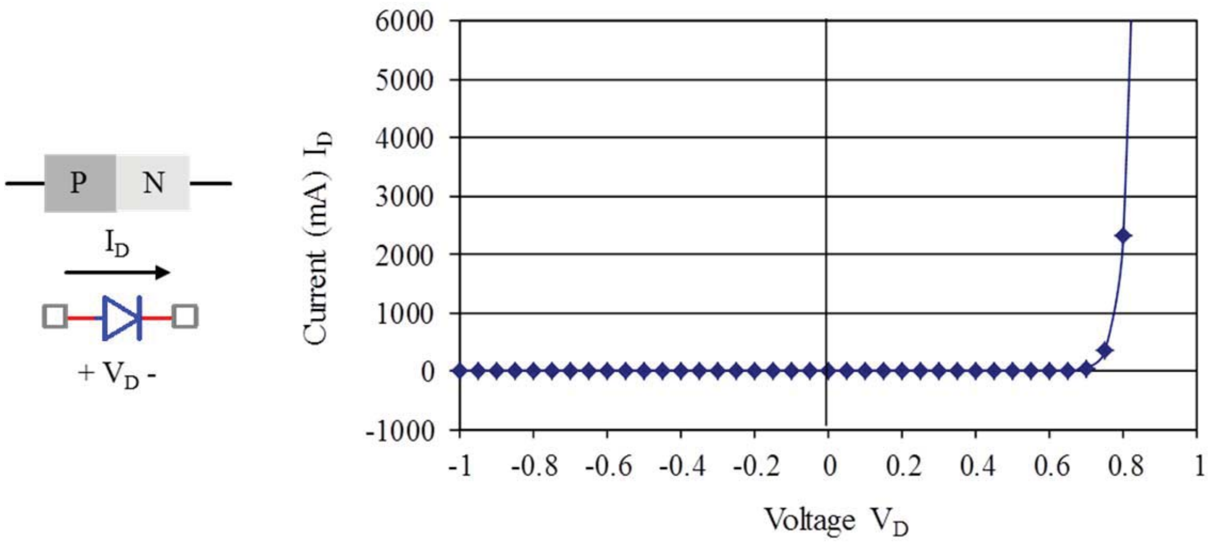


Figure 8.6: Example diode I-V curve

The I-V characteristics for a diode is defined by the following equation

$$I_D = I_S\left(e^{qV_D/KT} - 1\right) \tag{8.1}$$

In equation 8.1, q is the electron charge $\approx 1.6\times10^{-19}$ C, k is the Boltzmann constant $\approx 1.38\times10^{-23}\ \mathrm{m^2kg/s^2K}$, T is the temperature in Kelvin, and I_S is the saturation current which depends on the amount and type of doping and the physical geometry of the device. All the constants may be collapsed into a single variable V_T which is equal to about 26 mV. This reduces the diode I-V equation to the following:

$$I_D = I_S\left(e^{V_D/V_T} - 1\right) = I_S\left(e^{V_D/0.026} - 1\right) \tag{8.2}$$

We note that the I-V equation for a diode is not linear!

Examining Figure 8.6, we see that in reality this equation has two primary regimes. On regime is when V_D is less than some "turn-on" voltage and I_D is approximately zero. The other regime is when I_D is some positive value and V_D is approximately constant at the "turn-on" voltage. So we can in fact model this non-linear equation with a piece-wise linear model which has two regimes, "Off" and "On". This is called the "Large-Signal" diode model.

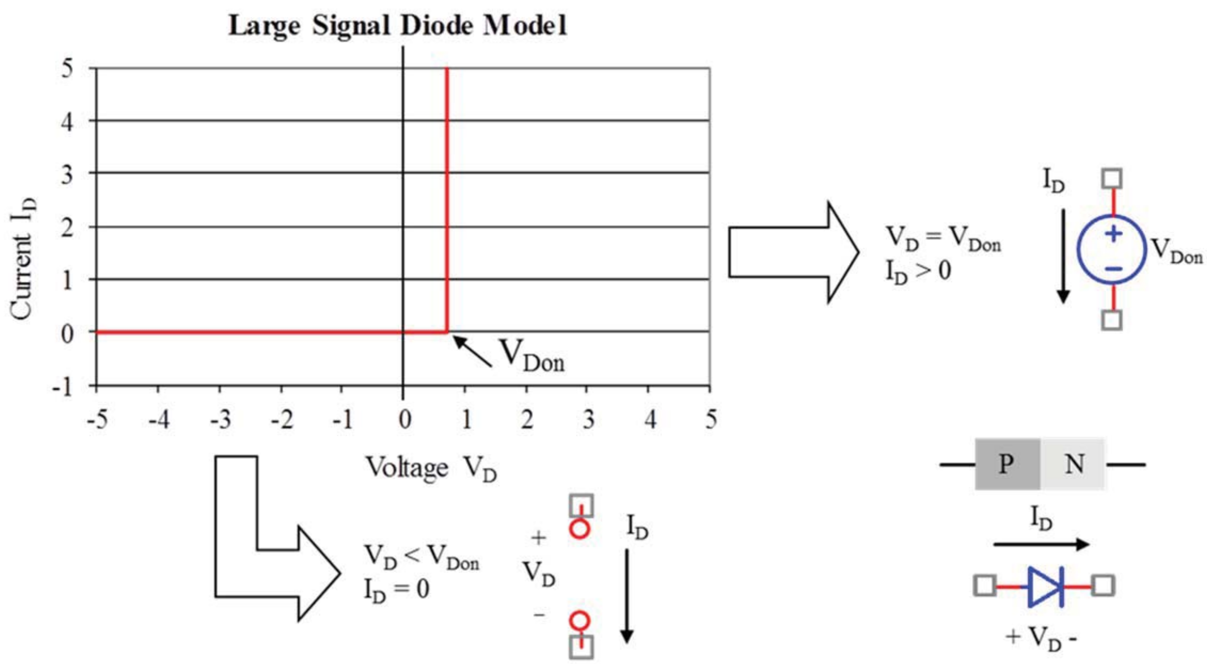


Figure 8.7: Large signal diode model showing the on and off regimes and their models

Figure 8.7 shows the Large Signal Diode Model. Again we note that this is a piece-wise linear model with two regimes. In other words, the model includes two segments which are themselves linear even though the entire model is not. The characteristics of these two regimes are as follows:

- Off regime: $V_D < V_{Don}$, $I_D = 0$
- On regime: $V_D = V_{Don}$, $I_D>0$

The value of V_{Don} depends on the specific diode and must be specified. For silicon diodes V_{Don} is approximately 0.7 V. For germanium diodes V_{Don} is approximately 0.3 V.

The model of the diode in the *Off* regime is an open circuit. The current through an open circuit is identically equal to zero. The model for the diode in the *On* regime is a DC voltage source with the value of V_{Don}. The voltage source will maintain a constant voltage across its terminals while it supplies whatever current the rest of the circuit requires.

8.5 Analysis of Diode Circuits

Consider the circuit shown in Figure 8.8. In order to analyze this circuit and determine all voltages and currents we first need to decide which diode model is most appropriate to use. In other words we first must guess which regime the diode is operating in, i.e whether the diode is on or off. The manner in which we make an educated guess is to see if the diode is possibly forward biased, in other words is V_D positive. If we believe that V_D is positive then we guess ON otherwise we guess OFF. Based on that guess we substitute our large signal model for the diode in the circuit and then solve the circuit using Kirchhoff's laws and Ohm's Laws. At this point we are not done! We need to check to see if our results are consistent with our initial guess.

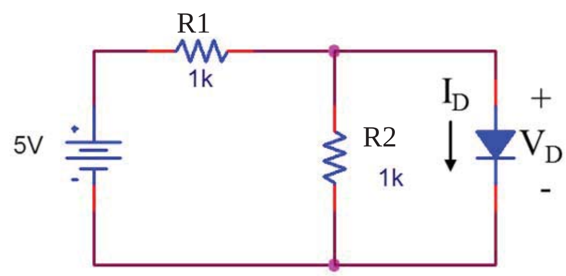


Figure 8.8: Example diode circuit

Let's apply these steps to the circuit shown in Figure 8.8. We will assume a silicon diode with V_{Don} = 0.7 V.

1. Assume either On or Off.
   To start let's assume OFF. Substitute into the circuit our diode large signal model for OFF, an open circuit (see Figure 8.9).
2. Solve for I_D and V_D using Kirchhoff's and Ohm's Law.
   I_D=0 (check!). Solve for V_D. V_D=2.5 V.
3. Check to see if our solution is consistent with original assumption.
   Is $V_D<V_{Don}$? No, V_D is greater than V_{Don}. Therefore our results are not consistent with the original assumption. This diode is not off.

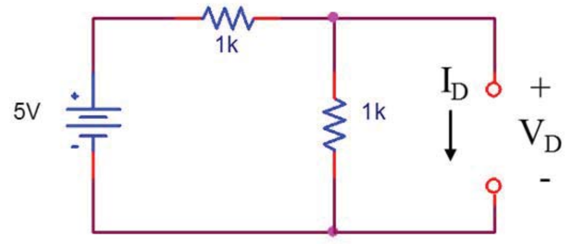


Figure 8.9: Example diode circuit assuming diode is OFF.

Let's start again only this time we will assume On.

1. Assume either On or Off.
   This time let's assume ON. Substitute into the circuit our diode large signal model for ON, a DC source equal to V_{Don} (see Figure 8.10).
2. Solve for I_D and V_D using Kirchhoff's and Ohm's Law.
   $V_D = 0.7$ V. Use Kirchhoff's Current Law to solve for I_D.
3.

$$\begin{aligned}\frac{5-V_D}{1000} &= \frac{V_D}{1000} + I_D \\ I_D &= \frac{4.3}{1000} - \frac{0.7}{1000} = 3.6\,\text{mA}\end{aligned} \tag{8.3}$$

4. Check to see if our solution is consistent with original assumption.
   Is $I_D>0$? Yes. Therefore our results are consistent with the original assumption. This diode is on.

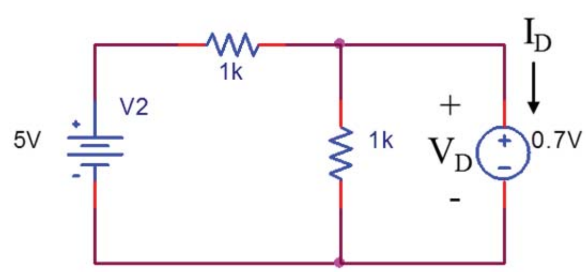


Figure 8.10: Example diode circuit assuming diode is ON.

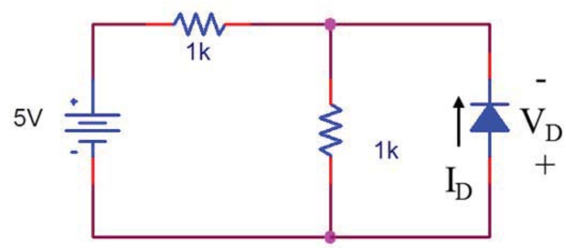
How small does the resistor parallel to the diode (R_2, see Figure 8.8) have to be in order for the diode to be turned off? The diode turns off when I_D is driven to zero. So we can determine the largest value for R_2 which will force the diode to operate at the point of transition from ON to OFF, i.e. $V_D = 0.7$ V and $I_D=0$ A. Under those conditions we can modify equation 8.3. to

$$
\begin{aligned}
\frac{5-V_D}{1000} &= \frac{V_D}{R_2} + I_D \\
\frac{4.3}{1000} &= \frac{0.7}{R_2} \\
R_2 &= 162.79\,\Omega
\end{aligned}
\tag{8.4}
$$

If $R_2 > 162.79\ \Omega$, then $V_D = V_{Don}$, $I_D > 0$, and the diode is on. If $R_2 < 162.79\ \Omega$, then $I_D = 0$, $V_D < V_{Don}$, and the diode is off.

THINK ABOUT IT

Consider the circuit shown. Is the diode on or off? What is V_D and I_D?



Consider the circuit with two diodes shown in Figure 8.11. With two diodes there are four possibilities: (i) both diodes are off; (ii) both diodes are on; (iii) D1 is on and D2 is off; and (iv) D1 is off and D2 is on. Again we make an educated guess based on whether or not the diodes look forward biased or reverse biased. Based on our guess we plug in the correct model for the diode and then solve the circuit for I_D and V_D of each diode. Finally we check to make sure that our solution is consistent with the original guess.

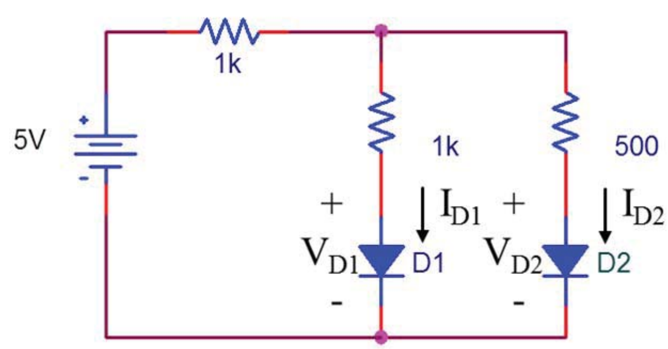


Figure 8.11: Example circuit with two diodes

1. Both diodes look forward biased so we will guess that they are both on. This means that $V_{D1} = V_{D2} = 0.7$ V.
2. We need to solve for the currents through the diodes. We can do this by applying KCL to the node connecting all the resistors:

$$
\begin{aligned}
\frac{5-V_A}{1000} &= \frac{V_A-0.7}{1000}+\frac{V_A-0.7}{500} \\
5-V_A &= V_A-0.7+2(V_A-0.7) \\
V_A &= 1.775
\end{aligned}
\tag{8.5}
$$

Next we solve for the currents:

$$
I_{D1}=\frac{V_A-0.7}{1000}=1.075\,\text{mA} \tag{8.6}
$$

and

$$
I_{D2}=\frac{V_A-0.7}{500}=2.15\,\text{mA} \tag{8.7}
$$

3. Finally we check that our solution is consistent with our original assumption (diodes on). If the diodes are on the $I_D>0$, which is what we determined. So we are consistent, the diodes are on.

THINK ABOUT IT
Consider the circuit shown in Figure 8.11, to what value would you need to change the 500 Ω resistor in order to turn D1 off.

8.6 Rectifier Circuits

As we have previously seen diodes can be used in circuits to act as switches. In other words, allowing current to flow only if the polarity of the applied voltage is correct. This characteristic is used in rectifier circuits. A rectifier circuit converts alternating current to direct current, AC to DC. The circuit shown in Figure 8.12 is a half-wave rectifier. Let's analyze how this circuit works.

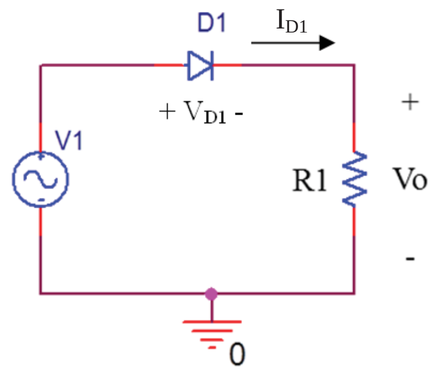
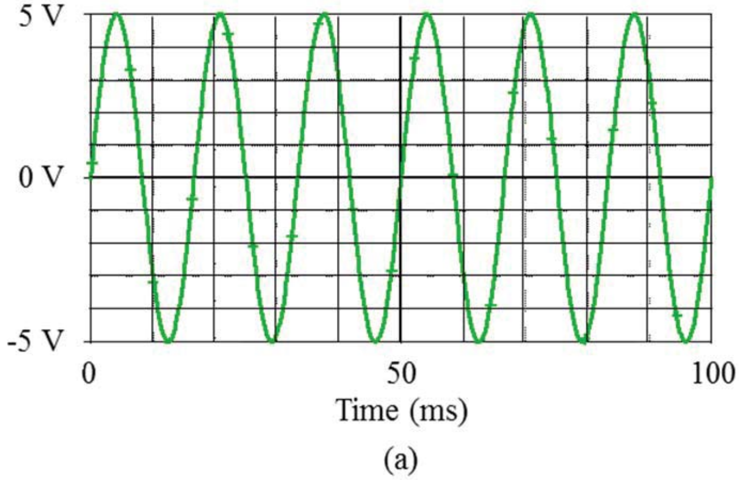


Figure 8.12: Half-wave rectifier circuit.

The voltage source, V1, is an AC voltage source. It produces a sine wave (see Figure 8.13(a)), similar to what comes out of a wall socket. During the positive half of the cycle the diode is forward biased and is therefore ON, i.e. $V_{D1}=V_{Don}$, $I_{D1}>0$. The current passes through the 1 kΩ resistor producing a voltage, Vo, that depends on the amount of current, i.e. $Vo=I_{D1}R_1$. Make note that V1 must be greater than the V_{Don} of the diode before current will flow. During the negative half of the cycle the diode is reverse biased and is therefore OFF, i.e. $V_{D1}<V_{Don}$, $I_{D1}=0$. Since no current exists the voltage across the resistor is zero, $Vo=I_{D1}R_1=0$. Figure 8.13(b) shows the voltage across the 1 kΩ resistor, Vo. Notice that we have converted a voltage that had both a positive and negative swing to a voltage that is always positive.



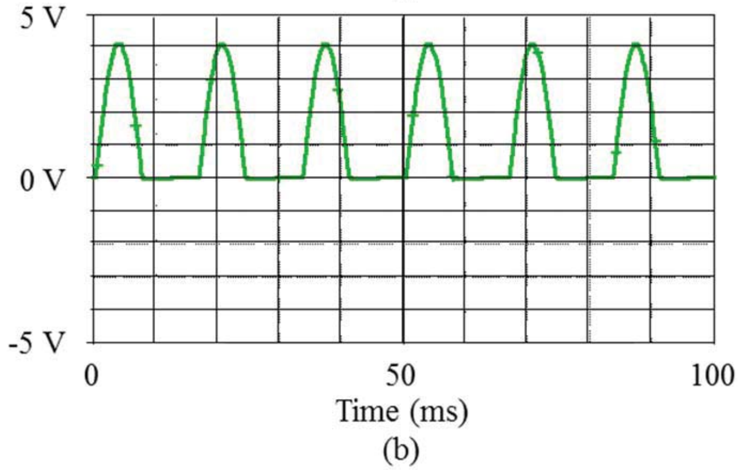


Figure 8.13: Voltage waveforms for (a) V1 and (b) Vo for the half-wave rectifier shown in Figure 8.12

Figure 8.13(b) does not yet look like a constant DC voltage. In order to *smooth* out the voltage that the resistor, or load, sees I need to add a capacitor as shown in Figure 8.14. You will learn about how capacitors work in EE203. For now we can think of the capacitor as a charge bucket. In other words when the diode is conducting current, some of the charge goes through the resistor and some of the charge goes into the capacitor and is stored there. When the diode stops conducting, the charge that was stored in the capacitor flows through the resistor in order to try to maintain the same voltage across the capacitor. The end result is an almost constant voltage across the resistor and capacitor. There is some *ripple* that we can define as ΔV which is the peak-to-peak voltage. Note that ΔV depends on the frequency of the original waveform, the value of the capacitor, and how much current the resistor requires to maintain the voltage.

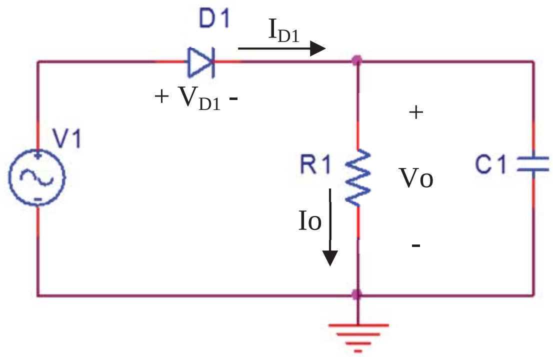


Figure 8.14: Half-wave rectifier with capacitor

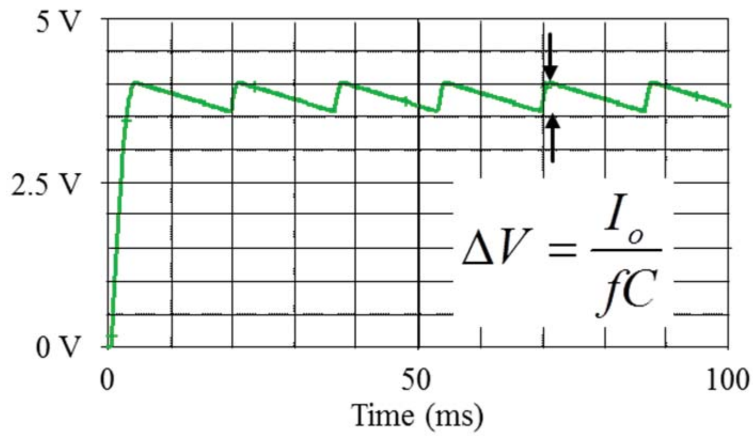


Figure 8.15: Voltage waveform for Vo of half-wave rectifier circuit shown in Figure 8.14.

The circuit shown in Figure 8.16 is a full-wave rectifier. Again this circuit converts an AC waveform into a DC waveform. Let's analyze how this circuit works. We are interested in knowing what the voltage, Vo, is across the *load*, R2. During the positive half cycle, D4 and D3 are forward biased and therefore ON. D5 and D2 are reverse biased and therefore OFF. The current then passes in a clockwise direction through the loop that includes V2, D4, R2, D3 and back to V2. The positive current direction is in passive sign convention with Vo and therefore Vo will be positive. During the negative half cycle of V2, D5 and D2 are forward biased and therefore ON, while D4 and D3 are reversed biased and therefore OFF. The current now passes in a counter-clockwise direction from V2 through D5, R2, D2, and back to V2. Note that the positive current direction through R2 is the same as previously. Therefore Vo is still positive. Figure 8.17(a) shows the sinusoidal voltage wave, V2. Figure 8.17(b) shows the rectified waveform, Vo. Note that the rectified waveform, Vo, is always positive. Again, if I place a capacitor across R2, I can smooth out the voltage across the resistor. The ripple voltage for a full-wave rectifier is given by:

$$\Delta V = \frac{Io}{2fC} \tag{8.8}$$

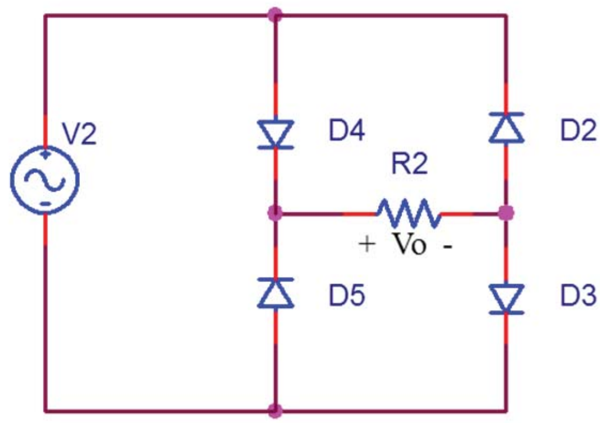


Figure 8.16: Full-wave rectifier circuit

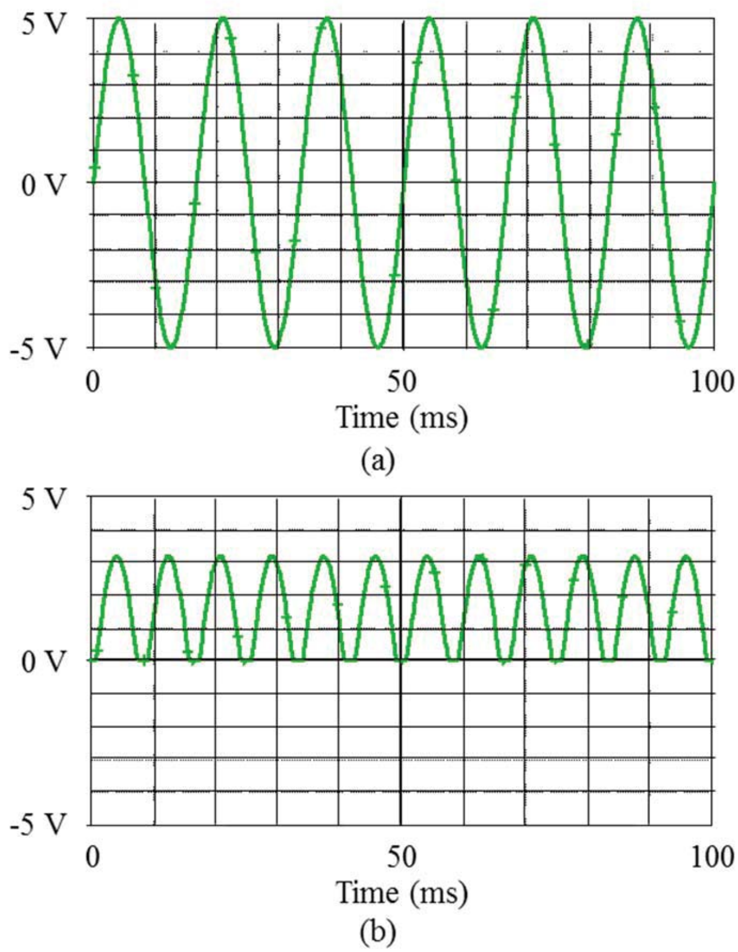


Figure 8.17: Voltage waveforms for (a) V2 and (b) Vo for the full-wave rectifier shown in Figure 8.16.

8.7 Zener Diodes

Zener diodes are a special type of diode which allows current to flow in the forward direction in the same manner as a regular diode, but also permits current to flow in the reverse direction if the voltage is above some *breakdown* voltage. In other words, instead of just two regimes ON and OFF, a Zener diode has three regimes, ON, OFF, and BREAKDOWN.

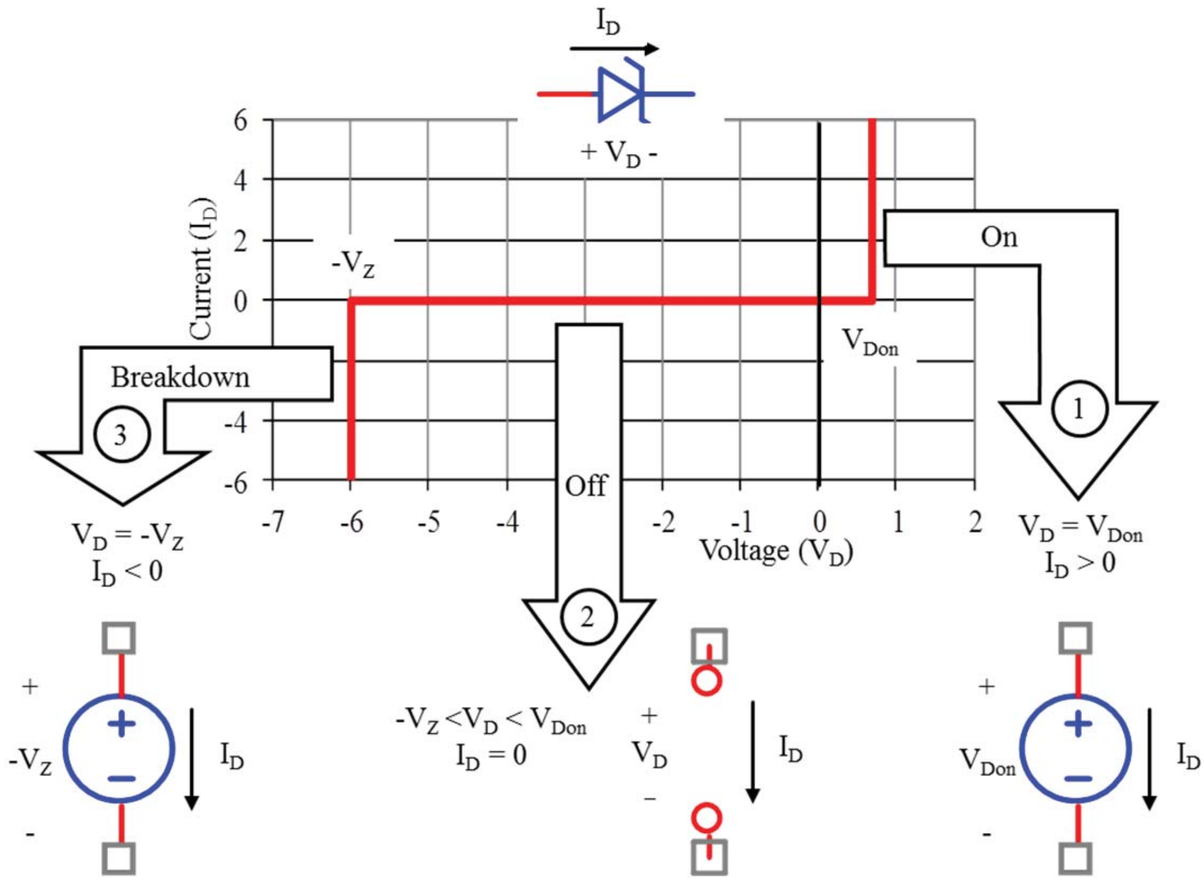


Figure 8.18: Large signal model of a Zener diode showing the (1) on, (2) off, and (3) breakdown regimes and their models

Figure 8.18 shows the I-V curve for an example large signal model of a Zener diode. In the (1) ON regime, the Zener diode operates in the same fashion as a regular diode, i.e. $V_D=V_{Don}$ and $I_D>0$. In this regime the model for the diode is a DC voltage source with a value of V_{Don}. In the (2) OFF regime, the Zener diode also operates in the same fashion as a regular diode, i.e. $I_D=0$. However, the OFF regime is limited to when V_D is greater than $-V_Z$ and less than V_{Don}, $-V_Z<V_D<V_{Don}$. The model for the Zener diode in this regime is an open circuit. For the (3) BREAKDOWN regime, the Zener diode conduct current in the negative direction, i.e. $I_D<0$ and the voltage across the Zener is $-V_Z$. In this regime we use another DC voltage source where $V_D= -V_Z$.

Let's apply our knowledge of Zener diodes to analyze the circuit shown in Figure 8.19. For this diode we will assume that $V_{Don} = 0.7$ V and $V_Z = 9$ V. The first thing we need to decide is which regime is the diode operating in: ON, OFF, or Breakdown? This diode appears to be forward biased so ON is a good choice. Under that assumption $V_D=V_{Don}=0.7$ V and we need to check that $I_D>0$.

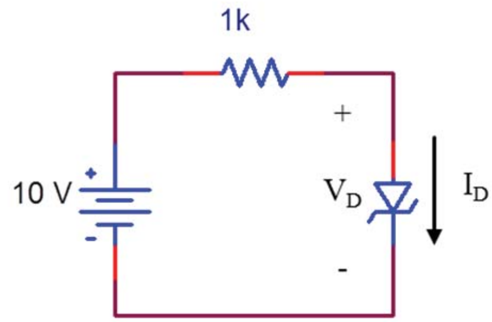


Figure 8.19: Example zener diode circuit

I_D is determined by applying Ohm's Law to the 1kΩ resistor.

$$I_D = \frac{10 - V_D}{R} = \frac{10 - 0.7}{1000} = 9.3\,\text{mA} \tag{8.9}$$

$I_D > 0$, which is consistent with our original assumption so the diode is ON.

Consider the circuit shown in Figure 8.20. Again, we will assume that $V_{Don} = 0.7$ V and $V_Z = 9$ V. Pay attention to the direction of I_D and the polarity of V_D. The direction of I_D and the polarity of V_D are pre-defined by the diode I-V characteristics (see Figure 8.18). This diode appears to be reverse biased so perhaps the diode is operating in the OFF regime. Under this assumption $I_D=0$ and we need to check that $-9<V_D<0.7$.

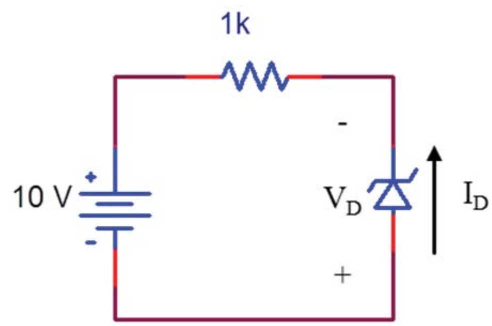


Figure 8.20: Example zener diode circuit

V_D is determined by applying KVL around the loop.

$$-10 - I_D(1000) - V_D = 0 \tag{8.10}$$

Remember that under our assumption that the diode is off $I_D=0$. Therefore solving equation 8.10 we find that

$$V_D = -10\,\text{V} \tag{8.11}$$

The value for VD is *not* consistent with our original assumption, i.e. $-9<V_D<0.7$, therefore our original assumption that the diode is operating in the OFF regime is incorrect. We need to try again.

This time let's assume that the Zener diode in the circuit shown in Figure 8.20 is operating in Breakdown. In Breakdown we model the Zener with a DC voltage source with a value of $-V_Z$ which in this case is -9 V. We need to check that $I_D<0$. Our KVL equation 8.10 is still valid under this new assumption so we can use that equation to solve for I_D.

$$-10 - I_D(1000) - (-9) = 0 \tag{8.12}$$

$$I_D = \frac{9-10}{1000} = -1\,\text{mA} \tag{8.13}$$

ID<0, which is consistent with our assumption that the diode is in breakdown.

Zener diodes are often used in their breakdown regime which provides voltage regulation. Operating in breakdown *regulates* the voltage across the load, in this case R_L (see Figure 8.21). In other words no matter what fluctuations or ripple occur on the 10 V supply, the voltage V_L will be a constant so long as the Zener diode remains in breakdown.

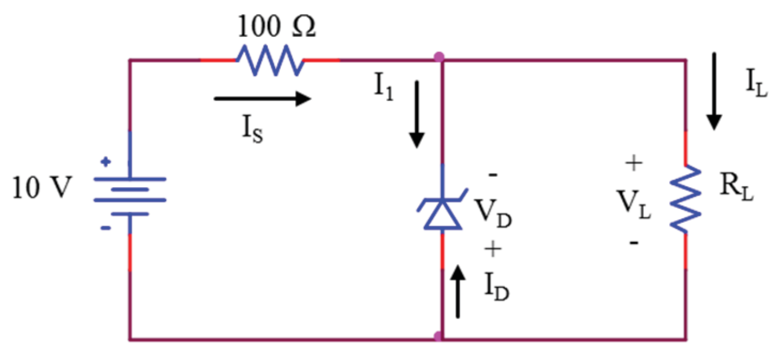


Figure 8.21: Example regulator circuit

Consider the circuit shown in Figure 8.21. We will assume that $V_{Don} = 0.7$ V and $V_Z = 9$ V. The Zener diode is reverse biased and so we can assume that it is operating in Breakdown. Under that assumption $V_D=-V_Z=-9$V and $I_D<0$. We can now solve for all voltages and currents in this circuit:

$$V_L = -V_D = 9\text{ V} \tag{8.14}$$

$$I_S = \frac{10-9}{100} = 10\,\text{mA} \tag{8.15}$$

Let's assume that $R_L = 1$ kΩ, then

$$I_L = \frac{9}{1000} = 9\,\text{mA} \tag{8.16}$$

Finally,

$$\begin{aligned} I_1 &= I_S - I_L = 1\,\text{mA} \\ I_D &= -I_1 = -1\,\text{mA} \end{aligned} \tag{8.17}$$

$I_D<0$, which is consistent with our assumption that the diode is in breakdown.

What is the minimum value for R_L such that the Zener diode in the circuit shown in Figure 8.21 remains in breakdown? As R_L gets smaller, it will require more current, I_L, in order to maintain the constant voltage V_L. At some point the larger I_L will force I_1 (and therefore I_D) towards zero (see equation 8.17). There is one point on the Zener diode I-V curve where $I_D=0$ and $V_D=-V_Z$. The load resistance that forces the Zener diode to operate at this point is the smallest resistance for which the Zener will be in breakdown. If the resistor is smaller than this value then $V_D>-V_Z$. So the smallest load resistance for which the Zener remains in breakdown is determined as follows:

$$I_1 = I_S - I_L = 0 \Rightarrow I_L = I_S = 10\,\text{mA} \tag{8.18}$$

$$V_L = I_L R_L \Rightarrow R_L = \frac{9}{10\,\text{mA}} = 900\,\Omega \tag{8.19}$$

If $R_L<900$ then the Zener diode will no longer be in breakdown.

Chapter 9: Transistors

In this chapter you will be introduced to transistors. We will examine how the sandwich of P-type and N-type semiconductors create bipolar junction transistors. We will examine the transistor I-V characteristics and learn how the three terminal transistor has both input and output I-V characteristics which are made up of families of curves rather than a single curve. As with the diode, we will develop a piece-wise linear model for the BJT transistor that we will use to analyze common emitter circuits.

9.1 Key Concepts

- **Transistors** are 3-terminal semiconductor devices used to amplify and switch electronic signals and electrical power. They are the building blocks for all modern electronics.
- **P-N-P and N-P-N** represents a sandwich of three doped regions which forms a transistor. A small electric current applied to the center layer (base) acts as a switch that allows a much larger current to flow between the outer layers (emitter and collector).
- **Bipolar Junction Transistor (BJT)** can be modeled as a current controlled current source.

9.2 History

Walter Brattain (Figure 9.1, right) and John Bardeen (Figure 9.1, left) announced their discovery of the "point contact" transistor in December 1947. William Shockley (Figure 9.1, center) thought that he should be given co-credit for the discovery as it was based on work that he did. Bell labs lawyers disagreed and the patent was made in Brattain and Barden's names only. Angry, Shockley set out to develop a better transistor in secret. Early in 1948 only a few months after the announcement of the point contact transistor, Shockley conceived of a transistor that looked like a sandwich, with two layers of one type of semiconductor surrounding a second type. This set-up created the bipolar junction transistor or BJT. In 1956 all three received the Nobel Prize in Physics for the invention of the transistor.

Figure 9.1: John Bardeen, William Shockley and Walter Brattain

9.3 Physics of BJT Transistors

There are two main classes of transistors (see Figure 9.2), the Bipolar Junction Transistor, or BJT, and the Field-Effect Transistor, or FET. The two types are different in the way they operate. You will learn the specifics of each when you take EE333. In this class we will focus on BJT transistors, the most common of which is the 2N2222. FETs have the advantage of being easier to manufacture, are smaller, and dissipate less power than BJTs. They are therefore the transistor of choice for integrated circuits. We will describe their operation when we talk about digital circuits.

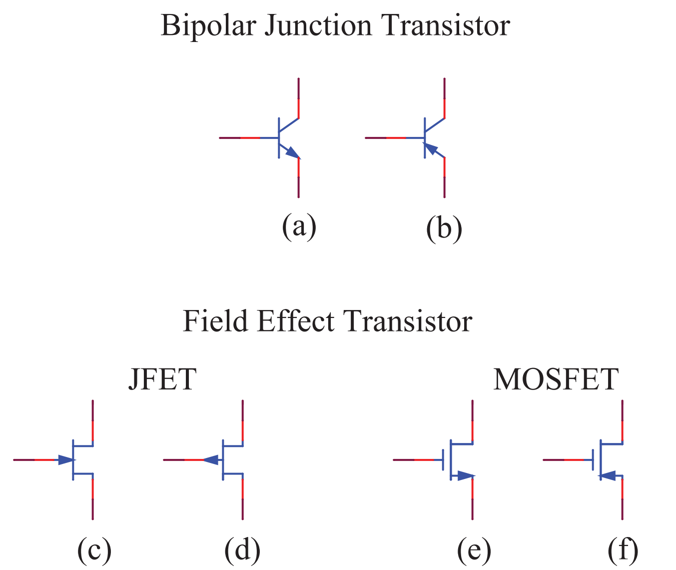


Figure 9.2: Circuit symbols for Bipolar Junction Transistors: (a) NPN, (b) PNP, and Field Effect Transistors: (c) N-type JFET, (d) P-type JFET, (e) N-channel enhancement mode MOSFET, (f) P-channel enhancement mode MOSFET.

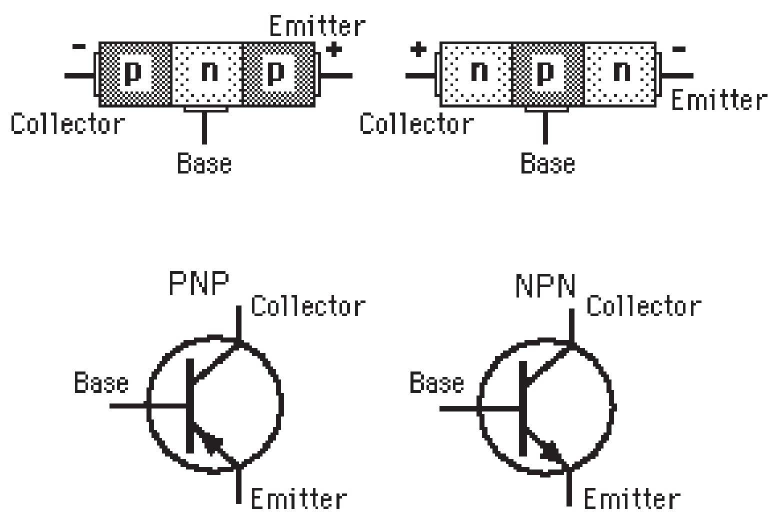


Figure 9.3: Bipolar junction transistors showing order of doped semiconductor with respect to the collector, base, and emitter of the transistor.

The bipolar junction transistor consists of three regions of doped semiconductors (see Figure 9.3). The order in which you sandwich the N- and P- type semiconductors determines whether you create an NPN type transistor or a PNP type transistor. The symbols for the NPN and PNP transistors are differentiated by the arrow in the emitter. If the arrow points from the base to the emitter it is an NPN transistor. If the arrow points from the emitter to the base it is a PNP transistor. Based on our previous discussion of PN junctions in the chapter 7, you might guess that the base-emitter junction of the transistor might act like a diode with the positive current direction being indicated by the emitter arrow. We will see in fact that the transistor can be modeled as a current amplifier when the base-emitter diode junction is "ON". A small current in the center or base region can be used to control a larger current flowing between the end regions (emitter and collector).

Shown in Figure 9.4, is an NPN transistor. The P-type material in the center is called the base and is narrower than the surrounding N-type emitter and collector regions. The emitter is heavily doped so as to be able to inject charged carriers into the base. Remember the current direction is opposite the direction of the electrons. The base is lightly doped and passes most of the injected charges to the collector. The collector region is the largest so it can dissipate heat.

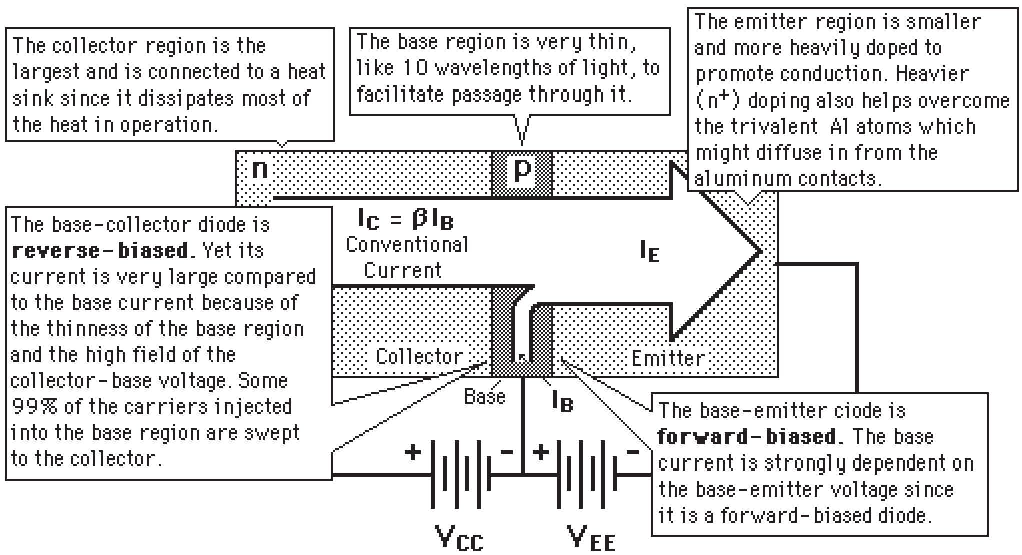


Figure 9.4: NPN bipolar junction transistor

As we will see shortly, the BJT transistor operation is characterized by three regions. For the NPN transistor these regions are defined as:

- **CUTOFF**: $V_{BE} < V_{BEon}$, i.e. the base emitter diode junction is OFF; $I_B = 0$; $I_C = 0$; and $V_{CE} > V_{CEsat}$.
- **ACTIVE**: $V_{BE} = V_{BEon}$, i.e. the base emitter diode junction is ON; $I_B > 0$; $I_C = \beta I_B$; and $V_{CE} > V_{CEsat}$.
- **SATURATION:** $V_{BE} = V_{BEon}$; $I_B > 0$; $I_C = I_{Csat}$; and $V_{CE} = V_{CEsat}$.

Consider Figure 9.5 showing current directions and voltage polarities for both an NPN and PNP transistor. Remember that KCL says that the sum of currents entering a node or region equals the sum of the currents exiting a note or region. KVL says that the sum of voltages around a closed loop must equal zero. Therefore:

$$I_B + I_C = I_E \qquad \text{(x.x)}$$

$$V_{CB} + V_{BE} - V_{CE} = 0 \qquad \text{(x.x)}$$

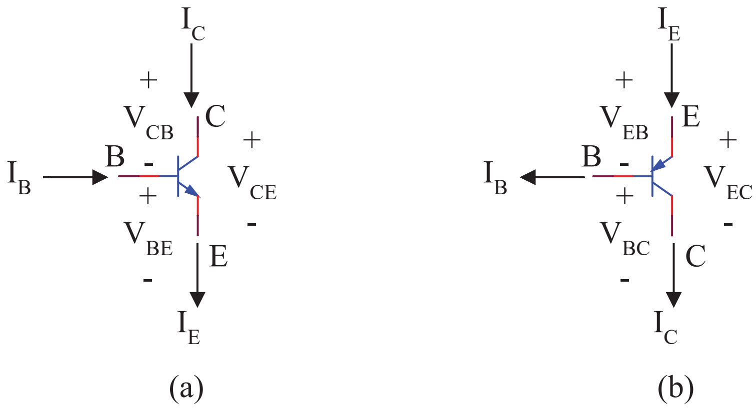


Figure 9.5: (a) NPN and (b) PNP transistors showing current directions and voltage polarities.

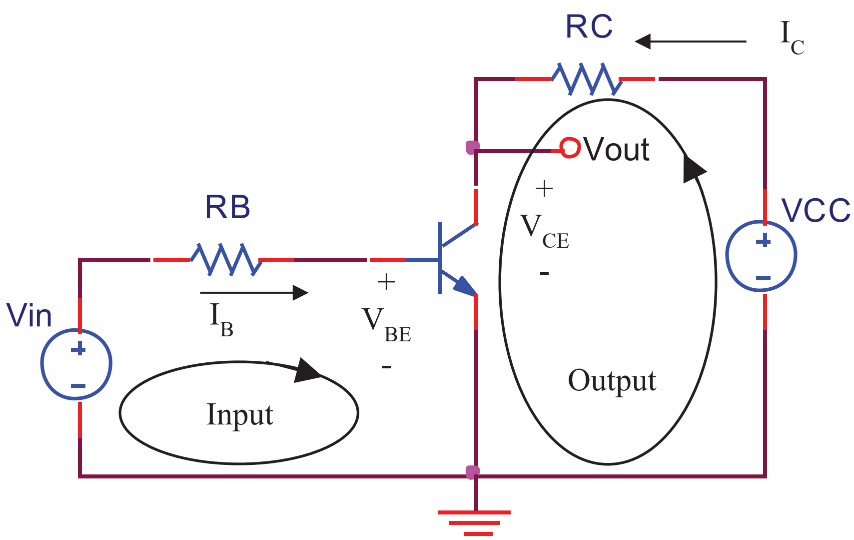


Figure 9.6: Common emitter circuit

9.4 BJT CE I-V Characteristics

Consider the circuit shown in Figure 9.6. In this circuit you can see two loops, an input loop and an output loop. The transistor in this circuit is configured with a common emitter connection (CE) which means that the emitter leg of the transistor is common to both the input and output loops. We can write KVL around both loops.

(I would prefer figure 9.6 goes here!!!)

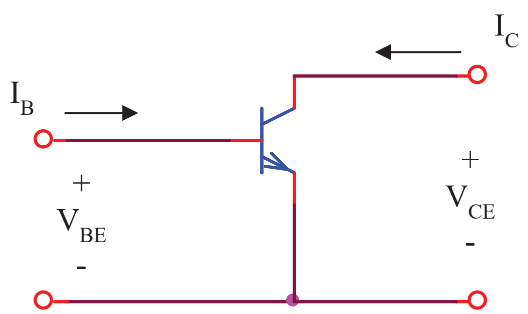
Going clockwise around the input loop we get:

$$-V_{in} + I_B R_B + V_{BE} = 0 \Rightarrow V_{in} = I_B R_B + V_{BE} \qquad \text{(x.x)}$$

Going counter clockwise around the output loop we get:

$$-V_{CC} + I_C R_C + V_{CE} = 0 \Rightarrow V_{out} = V_{CE} = V_{CC} - I_C R_C \qquad \text{(x.x)}$$

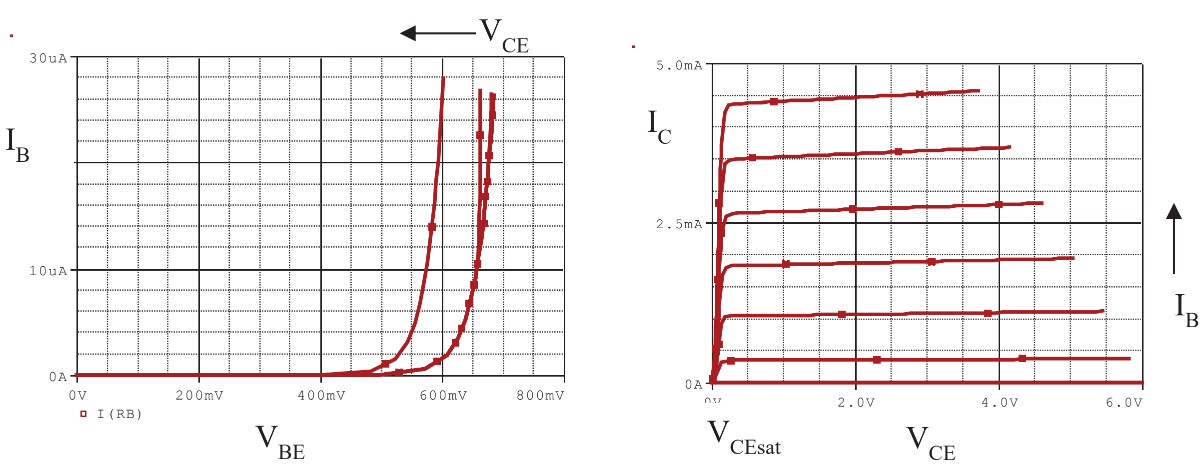
Because the transistor is a three terminal device there is more than one I-V curve that is of interest. In the common emitter configuration we can define an input I-V curve which shows the relationship of I_B as a function of V_{BE}, and an output I-V curve which shows the relationship of I_C as a function of V_{CE} (see Figure 9.7). However, for each I-V curve (input or output) there is not one curve but a family of curves since the input and output are coupled through the transistor operations. Therefore the I-V curve gives I_B not just as a function of V_{BE}, but also as a function of V_{CE}. The output I-V curve gives I_C not just as a function of V_{CE}, but also as a function of I_B.



input I-V output I-V

Figure 9.7: NPN Transistor showing input and output I-V definitions

Figure 9.8(a) shows the input I-V curves, i.e. I_B as a function of V_{BE} and V_{CE} for the NPN transistor. Note the similarity of the input I-V characteristics with a diode I-V curve. For the most part the input I-V characteristics are only slightly influenced by V_{CE}. However, as V_{CE} approaches V_{CEsat}, the input I-V curve will be forced to lower voltages. For most situations we can model the input I-V characteristics as a single curve similar to our diode mode.



(a) (b)

Figure 9.8: (a) Input and (b) output I-V curves.

The output I-V characteristics (Figure 9.8(b)) are more complicated. The graph here shows a family of curves that look suspiciously like current sources, i.e. they are approximately horizontal lines. In this class we will model them as exactly horizontal lines, however you will note that they are not really horizontal. The reasons for this will be explained in EE333. On which curve a specific circuit resides is strongly dependent on the value of I_B. As I_B gets larger I_C gets larger. In effect a BJT transistor can be thought of as a current controlled current source or as a current amplifier.

9.5 BJT CE Circuit Analysis

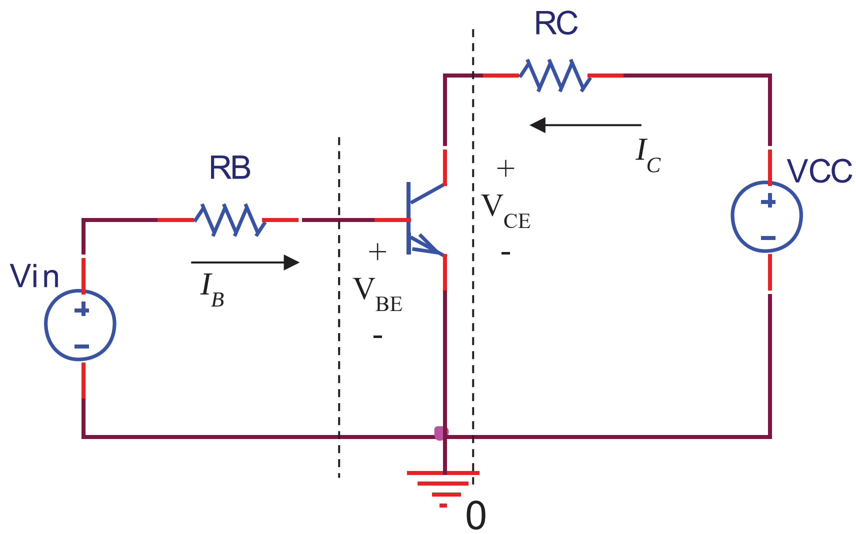


Figure 9.9: Common Emitter (CE) circuit analysis

Consider the common emitter circuit shown in Figure 9.9. We start our analysis of this circuit by writing Kirchhoff's Voltage Law around the input loop and solving for I_B.

$$I_B = \frac{V_{in} - V_{BE}}{R_B} \qquad \text{(x.x)}$$

We need to remember that the input I-V characteristic is that of a diode and can be approximated by one curve which we will model with two regimes CUTOFF and ON (Figure 9.10). In other words, if $V_{BE}<V_{BEon}$ then the transistor is in CUTOFF and $I_B=0$. When the base to emitter junction is forward biased enough then $V_{BE}=V_{BEon}$ and I_B will be greater than zero. The transistor is now ON. I_B must always be positive, therefore, V_{in} must be V_{BEon} or greater in order to turn on the transistor.

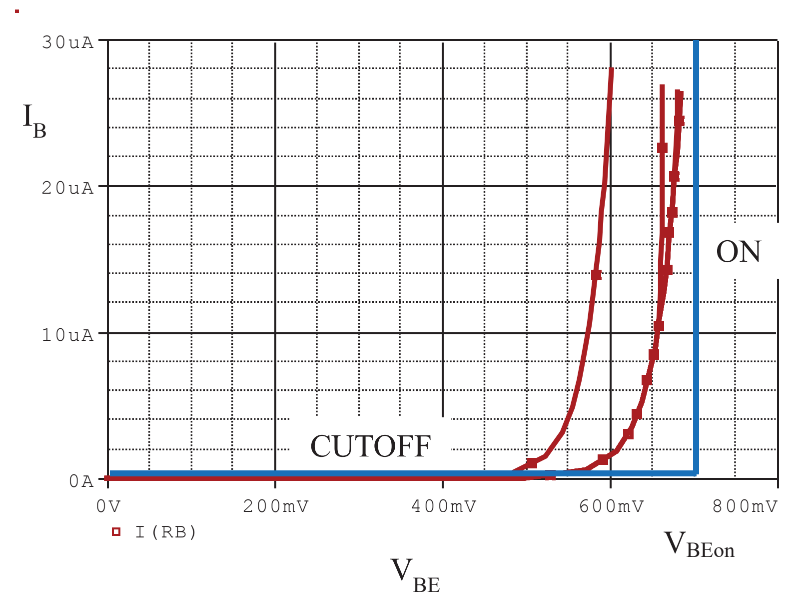


Figure 9.10: Input I-V curve showing piece-wise linear model

We continue our analysis of the common emitter circuit by writing Kirchhoff's Voltage Law around the output loop. Remember that I_C is a function of I_B and V_{CE}. The dependence on I_B cannot be eliminated. We can use the load like technique where we plot the output load (R_C and V_{CC}) I-V characteristics on top of the transistor output I-V characteristics (see Figure 9.11). Note that the solution strongly depends on I_B.

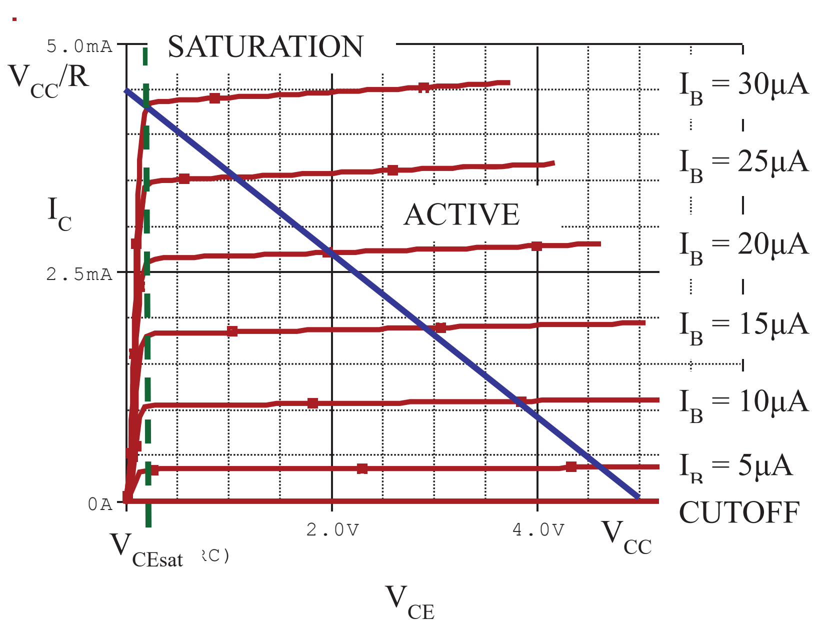


Figure 9.11: Output I-V curve showing R_C-V_{CC} load line.

We observe that the ratio of I_C/I_B is a constant, in this case equals 120. The ratio of I_C to I_B is called β. β defines the constant ratio of I_C to I_B in the active region and is a function of the transistor. Note that this relationship also holds when the transistor is in CUTOFF. As I_B

increases the operating point on the load line goes up, i.e. I_C increases and V_{CE} decreases. When V_{CE} is reduced to V_{CEsat}, the transistor is said to be in SATURATION and the ratio of I_C to I_B no longer holds. From the load line the maximum I_C is:

$$I_{C\max} = \frac{V_{CC} - V_{CEsat}}{R_C} \tag{x.x}$$

To summarize (see Table 9.1), the transistor has two sets of I-V curves that define its input and output I-V characteristics. These define three regions of operation for the transistor in its common emitter configuration: CUTOFF, ACTIVE, and SATURATION. The input I-V characteristic determines whether the transistor is in CUTOFF or ACTIVE in the same way a diode I-V characteristic determines whether the diode is ON or OFF. In CUTOFF, $V_{BE}<V_{BEon}$ and $I_B = 0$. In the ACTIVE region $V_{BE}=V_{BEon}$ and $I_B>0$. The input characteristics when the transistor is in SATURATION is the same as when the transistor is operating in the ACTIVE region, i.e. $V_{BE}=V_{BEon}$ and $I_B>0$.

The output I-V characteristic determines whether the common emitter circuit is in the ACTIVE or SATURATION regions. If ACTIVE, then $I_C = \beta I_B$ and V_{CE} is defined by KVL, i.e. $V_{CE}=V_{CC}-I_C R_C$. If in SATURATION, then I_C is so large that V_{CE} is driven to equal V_{CEsat}. When that happens I_C cannot get any larger and $I_C=I_{Cmax}$. When the transistor is in CUTOFF all currents are zero so $I_C=0$. KVL around the output loop then gives $V_{CE} = V_{CC}$.

Table 9.1: Summary model equations for common emitter circuit shown in Figure 9.9.

	Input Characteristics	Output Characteristics
Cutoff	$V_{BE} < V_{BEon}$ $I_B = 0$	$I_C = 0$ $V_{CE} = V_{CC}$
Active	$V_{BE} = V_{BEon}$ $I_B > 0$	$I_C = \beta I_B$ $V_{CE} = V_{CC} - R_C I_C$
Saturation	$V_{BE} = V_{BEon}$ $I_B > 0$	$I_C = I_{C\max} = \frac{V_{CC} - V_{CEsat}}{R_C}$ $V_{CE} = V_{CEsat}$

9.6 BJT Small Signal Model

BJT transistors can operate in three different regimes, CUTOFF, ACTIVE, and SATURATION. We then need three different models to reflect these three regimes of operation (see Figure 9.12).

The transistor is in CUTOFF when $V_{BE}<V_{BEon}$ causing the base emitter diode to be off. Therefore all currents are zero. We model this state for the transistor in the same way we modeled the diode, i.e. as an open circuit (see Figure 9.12(a)).

In the ACTIVE region $V_{BE}=V_{BEon}$ and $I_C = \beta\ I_B$. Therefore in the active region we model the transistor base emitter junction with a DC voltage equal to V_{BEon} and the collector emitter junction with a current controlled current source $I_C = \beta I_B$ (see Figure 9.12(b)). Remember that the values of V_{BEon} and β are transistor dependent.

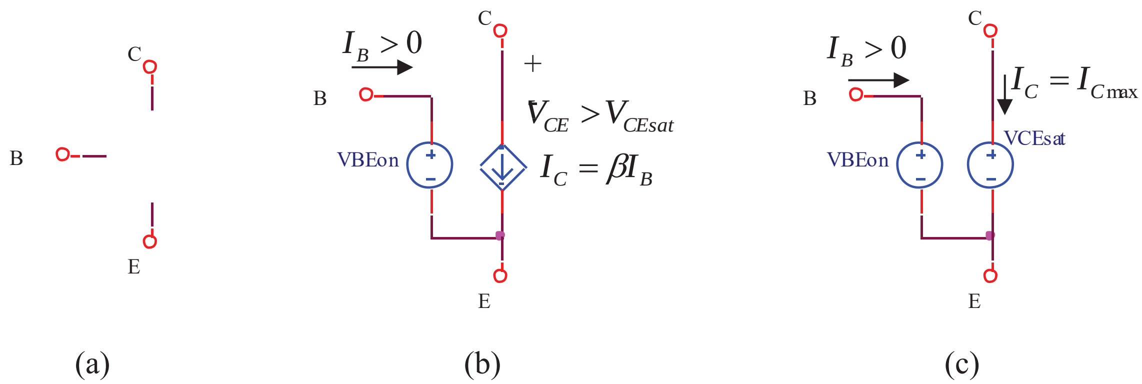


Figure 9.12: BJT Small Signal Models for transistor operating in (a) Cutoff, (b) Active, and (c) Saturation

For a transistor in SATURATION, I_C is no longer dependent on I_B, so the current controlled current source is no longer an appropriate model. However, we do know that $V_{CE} = V_{CEsat}$ when the transistor is in saturation. Therefore we model the transistor in saturation with a DC voltage equal to V_{BEon} for the base emitter junction and another DC voltage equal to V_{CEsat} for the collector-emitter junction (see Figure 9.12(c)).

Using the models shown in Figure 9.12 we can examine what the common emitter circuit looks like when the transistor is operating in each regime. For example, consider the common emitter circuit shown in Figure 9.13(a). The value of V_{in} determines whether the transistor in this circuit is in CUTOFF, in the ACTIVE region, or in SATURATION. We saw before that if $V_{in}<V_{BEon}$ then the transistor is in CUTOFF and all currents are zero (see Figure 9.13(c)). Looking at this circuit we can see that if all currents are zero then there will be no voltage drops across the R_B or R_C resistors. Therefore $V_{BE}=V_{in}$ and $V_{CE}=V_{CC}$. If V_{in} becomes larger than V_{BEon} the transistor goes into the ACTIVE region. In this region the base emitter junction is modeled by a DC voltage source and $I_B>0$ (see Figure 9.13(b)). The collector emitter junction is modeled with a current controlled current source $I_C = \beta I_B$ and $V_{CC}>V_{CE}>V_{CEsat}$. If V_{in} increases more the transistor will go into SATURATION. In saturation the collector emitter junction is modeled with a DC voltage equal to V_{CEsat} (see Figure 9.13(d)).

Figure 9.14 shows graphically what happens to V_{BE}, I_B, V_{CE}, and I_C as V_{in} increases. The input characteristics are shown in the upper two plots. When $V_{in} = 0$ both V_{BE} and I_B equal zero. As V_{in} increases I_B remains zero since the transistor is still in CUTOFF regime. V_{BE} increases linearly with V_{in} until $V_{BE}=V_{BEon}$. If V_{in} increases more, V_{BE} remains equal to V_{BEon} while I_B increases with V_{in} throughout the ACTIVE and SATURATED regimes. The output characteristics are shown in the bottom two plots. Again when $V_{in} = 0$, $I_C = 0$ and remains zero while the transistor is in CUTOFF. Since $I_C=0$, KVL gives us that $V_{CE}=V_{CC}$ (in other words there is no voltage drop across the collector resistor). When the transistor is in the ACTIVE regime, I_C increases with I_B. Increasing I_C means increasing voltage drop across the collector resistor, so V_{CE} must decrease. At some point as V_{in} continues to increase V_{CE} reaches V_{CEsat}. The transistor is now in SATURATION. V_{CE}=V_{CEsat} and I_C=I_{Cmax}.

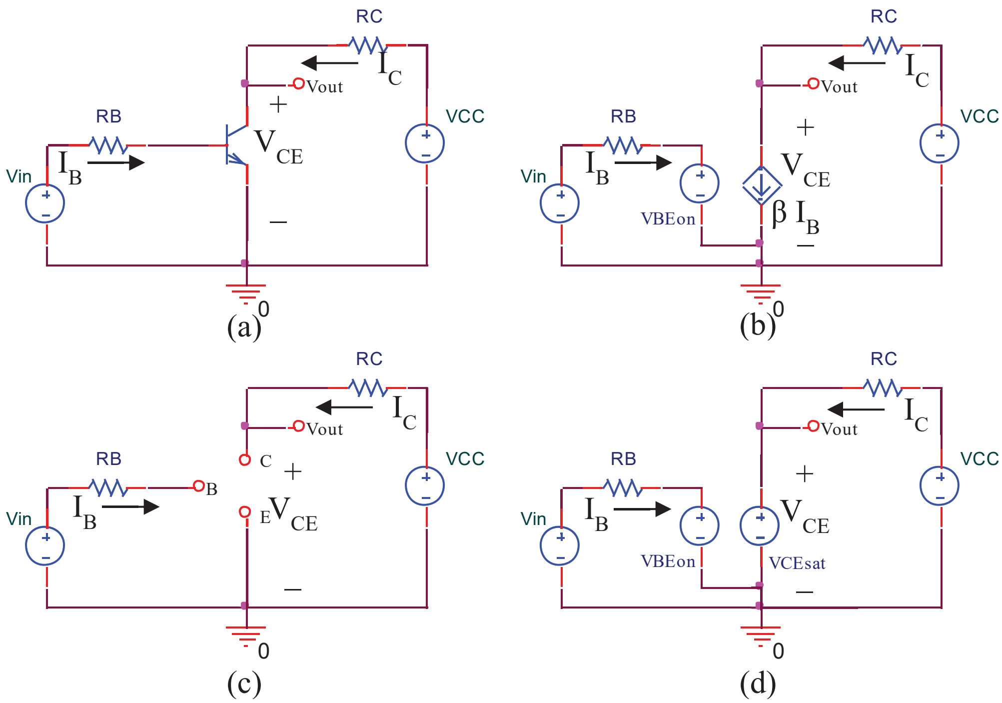


Figure 9.13: (a) Common emitter amplifier showing (b) active model, (c) cutoff model, and (d) saturation model.

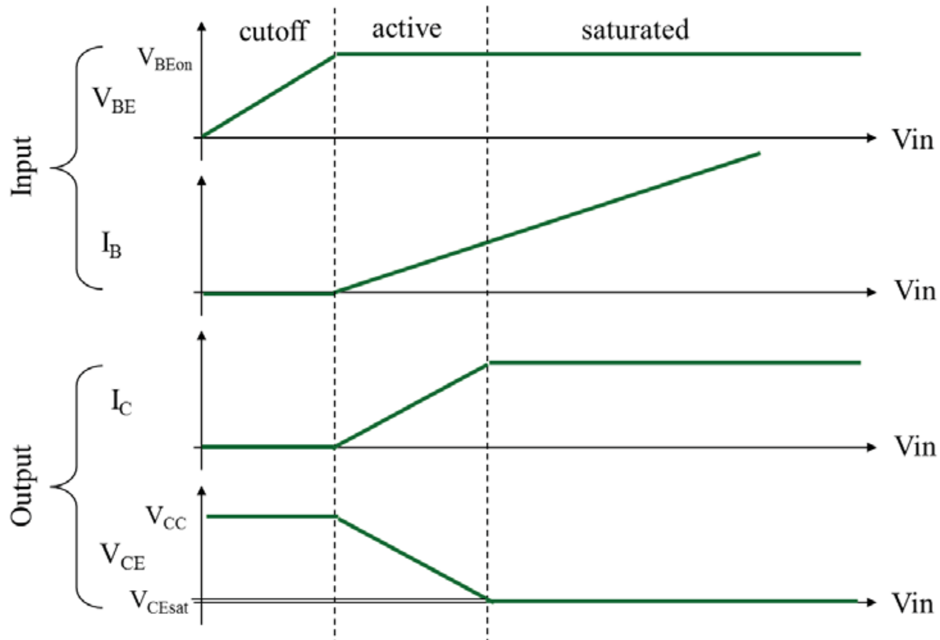


Figure 9.14: CE characteristics versus Vin for (a) input and (b) output loops.

9.7 BJT Power Dissipation

Consider the common emitter circuit shown in Figure 9.15. We are interested in knowing how much power is being dissipated by the transistor. We can approach this problem by applying conservation of energy. Conservation of energy says that the sum of the power dissipation in the input circuit, the transistor, and the output circuits will exactly equal zero. In other words the power dissipated in the transistor is the negative of the power dissipated in the input and output circuits.

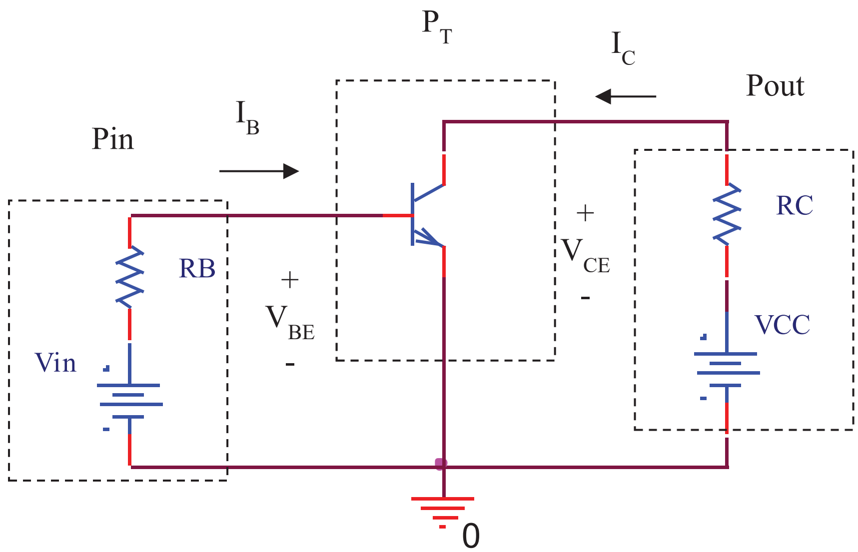


Figure 9.15: Common emitter circuit.

$$P_{in} + P_T + P_{out} = 0 \qquad \text{(x.x)}$$

Equation (x.x) may be re-written to be a function of I_B and V_{BE} (input circuit power) and I_C and V_{CE} (output circuit power) to find that the power dissipated by the transistor, P_T.

$$P_T = -P_{in} - P_{out} = V_{BE} I_B + V_{CE} I_C \qquad \text{(x.x)}$$

When the transistor is in cutoff, $V_{BE}=V_{in}$, $V_{CE} = V_{CC}$ and all currents are zero, therefore the power dissipated by the transistor in cutoff is also zero.

$$P_{Tcutoff} = V_{in} \cdot 0 + V_{CC} \cdot 0 = 0 \qquad \text{(x.x)}$$

In the active region, $V_{BE}=V_{BEon}$, $I_C = \beta\ I_B$, and $V_{CE} = V_{CC}- I_C\ R_C$ or substituting in for I_C, $V_{CE} = V_{CC}-\beta\ I_B\ R_C$. We can then write the power dissipated by the transistor in terms of the base current I_B.

$$P_{Tactive} = V_{BEon} I_B + \left(V_{CC} - \beta I_B R_C\right)\beta I_B \qquad \text{(x.x)}$$

The value of I_B for which the power dissipation in maximum is determined by setting the derivative of equation (x.x) with respect to I_B to zero and solving for I_B. Plugging the result back into the equation (x.x) gives maximum power dissipated in the active region.

$$P_{T\max} = \frac{(V_{BEon} + \beta V_{CC})^2}{4\beta^2 R_C} \quad \text{(x.x)}$$

Finally, in Saturation, $V_{BE}=V_{BEon}$, $V_{CE}=V_{CEsat}$, and $I_C=I_{Cmax}$. Therefore the power dissipated while in saturation is given by.

$$P_{Tsaturation} = V_{BEon}I_B + V_{CEsat}I_{C\max} = V_{BEon}I_B + \frac{V_{CEsat}(V_{CC} - V_{CEsat})}{R_C} \quad \text{(x.x)}$$

Figure 9.16 shows an example of a BJT transistor power dissipation curve with respect to I_B. When the transistor is in CUTOFF the power dissipated by the transistor is zero. In the ACTIVE regime the power dissipated by the transistor is quadratically dependent on I_B, starting at zero, increasing to a maximum when I_B is approximately half the distance to saturation and then decreasing to a minimum when the transistor transitions into SATURATION. In Saturation the power dissipated by the transistor is dominated by $I_{C\max}V_{CEsat}$ and is approximately constant. Transistors used in digital circuits typically operate in cutoff and saturation thereby minimizing the power dissipation in the digital circuit.

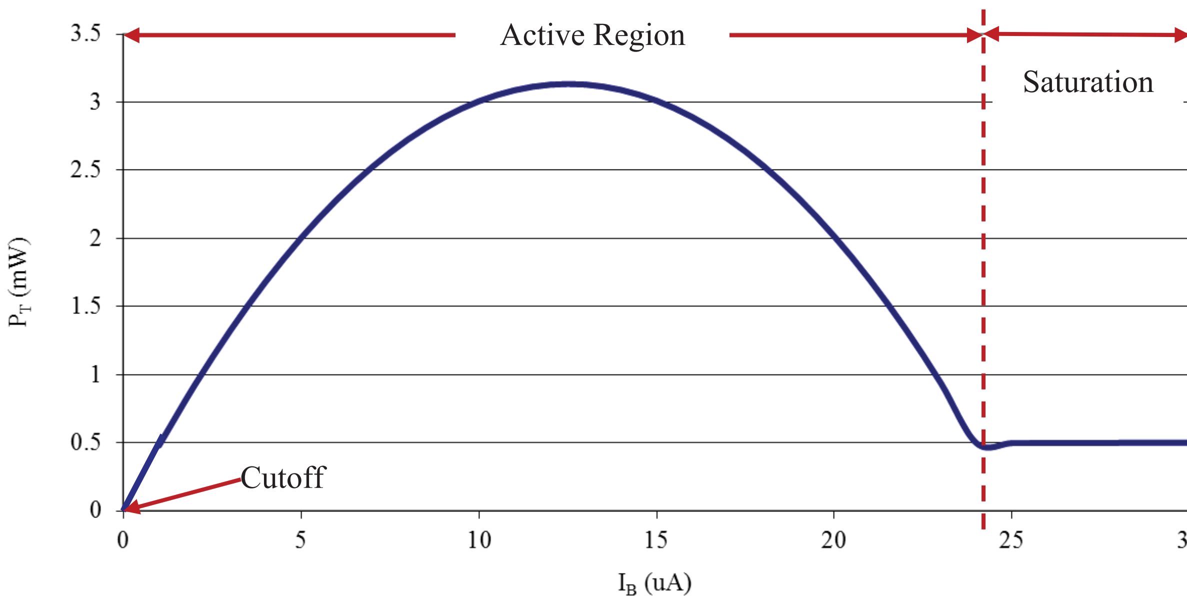


Figure 9.16: Example power dissipation curve for a BJT transistor.

9.8 BJT CC Circuit Analysis

The common emitter configuration is not the only way one can use a transistor. Figure 9.17 shows the transistor in a common collector or emitter follower configuration. Our approach to analyzing this circuit is the same as the common emitter circuit. In other words we write KVL around the input loop and the output loop. Going clockwise around the input loop gives

$$-V_{in} + V_{BE} + I_E R_E = 0 \Rightarrow I_E = \frac{V_{in} - V_{BE}}{R_E} \quad \text{(x.x)}$$

Again, in order for the transistor to "turn on", $V_{in}>V_{BEon}$. I_E must be greater than zero. Writing KVL around the output loop gives

$$-V_{CC}+I_C R_C+V_{CE}+I_E R_E=0 \qquad \text{(x.x)}$$

Equation (x.x) has two unknowns, I_C and V_{CE}. I_E was determined from equation (x.x). If the transistor is operating in the active region then $I_C=\beta I_B$. For this circuit I_B is unknown, however I_E is known. The relationship between I_E and I_C is determined by applying KCL around the transistor, so

$$I_E=I_C+I_B=I_C+\frac{I_C}{\beta}=\left(\frac{\beta+1}{\beta}\right)I_C=\alpha I_C \qquad \text{(x.x)}$$

The coefficient $\beta/(\beta+1)$ is called α and is defined as the ratio of I_E to I_C, just like β was defined as the ratio of I_C to I_B. So now the output loop equation can be reduced to

$$I_C=\frac{V_{CC}-V_{CE}}{R_C+\left(\frac{\beta+1}{\beta}\right)R_E} \qquad \text{(x.x)}$$

or

$$V_{CE}=V_{CC}-I_E\left(\left(\frac{\beta}{\beta+1}\right)R_C+R_E\right) \qquad \text{(x.x)}$$

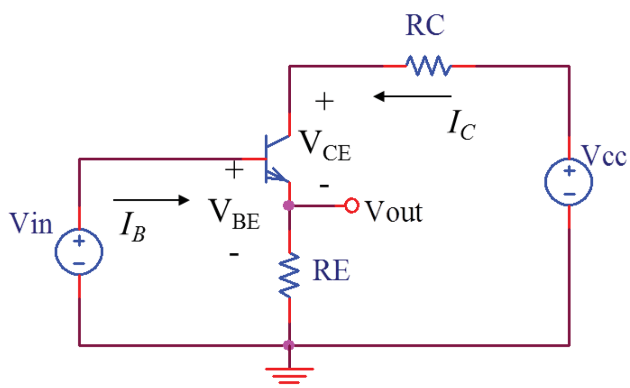


Figure 9.17: Common collector amplifier.

9.9 BJT CE Circuit Design

Consider the circuit shown in Figure 9.18. We wish to design this circuit such that the quiescent point $Q(I_C, V_{CE})$ of the transistor is $I_C = 1mA$ and $V_{CE} = 5\ V$, when $V_{in} = 1\ V$. The quiescent point, also known as the DC bias point or Q-point, defines the operating value of V_{CE} and I_C when the small signal input is zero. This definition is specifically applicable to using the common emitter as an amplifier. We can think of V_{in} as being represented by a DC bias voltage added to a sinusoidal signal. The sinusoidal signal is what is intended to be amplified. You will learn more about designing amplifiers in EE333. So in effect we need to choose R_B, R_E, and R_C such that when $Vin = 1\ V$: $I_C = 1\ mA$, and $V_{CE} = 5\ V$. These are the design parameters or requirements.

The approach to design is similar to what was previously done for analysis. We start with the KVL equations for the input and output loops.

$$V_{in} = I_B R_B + V_{BEon} + I_E R_E \quad \text{(x.x)}$$

and

$$V_{CC} = I_C R_C + V_{CE} + I_E R_E \quad \text{(x.x)}$$

The design requirements specify I_C = 1mA and V_{CE} = 5 V when V_{in} = 1 V, so we can plug in for those values. We also know that if we are in the ACTIVE regime that $I_B=I_C/\beta$ and $I_E=(\beta+1)/\beta\, I_C$. For this example we will assume that the transistor V_{BEon}=0.7 V and β=100. Substituting in for the known values gives

$$1 = \left(\frac{0.00001}{100}\right) R_B + 0.7 + \left(0.001\frac{101}{100}\right) R_E \quad \text{(x.x)}$$

and

$$10 = (0.001) R_C + 5 + \left(0.001\frac{101}{100}\right) R_E \quad \text{(x.x)}$$

Equations (x.x) and (x.x) have three unknowns. Solving for three unknowns requires a minimum of three equations. Previously the common emitter circuit didn't have an emitter resistor. The emitter resistor is typically added to increase stability at the cost of decreasing the gain. A convenient rule of thumb is to choose $R_B = \beta * R_E/10$. This now becomes the third equation and all resistors can be solved for. You will learn more about how the emitter resistor affects gain and stability in EE333.

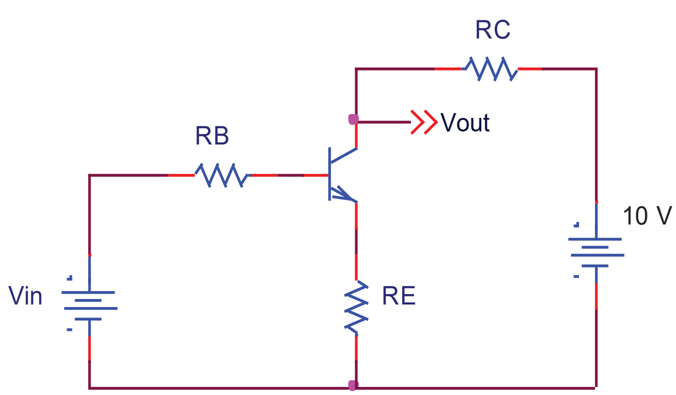


Figure 9.18: Common emitter amplifier with emitter resistor.

Chapter 10: Digital Circuits

In this chapter we will explore the creation of the computer age starting with George Boole's formulation of Boolean algebra and the resulting logic gates. We will specifically define Boolean Operations and Algebraic Identities and use those operations and identities to create truth tables and circuits that represent Boolean functions. Finally, we will explore how all Boolean functions may be implemented using a two-level AND-OR circuit; how this two-level circuit may be derived by determining the Minterms of the truth table; and how to use Karnaugh Maps to optimize or minimize the logic required to implement the Boolean function.

10.1 Key Concepts

- **Digital Circuits** are circuits that manipulate binary information
- **A Gate** is a circuit (made up of transistors, resistors, etc.) whose inputs and output are binary variables. A gate performs a specific Boolean (logical) operation.

10.2 History

George Boole is sometimes regarded as the founder of computer science, although computers were not yet invented. In 1854 he published "An Investigation of the Laws of Thought, on which are founded the mathematical theories of logic and probability". Boole did not regard logic as a branch of mathematics, as the title implies, but suggested that logic and algebraic symbols were similar. Boole's system was based on a binary approach processing only two objects, the yes-no, true-false, on-off, zero-one approach. Boole's work, as well as that of his progeny, were relatively obscure except among logicians, and seemed to have no practical use.

Figure 10.2: George Boole
1815-1864

Approximately seventy years after Boole's death, Claude Shannon discovered Boolean algebra while taking a philosophy class at the University of Michigan. Shannon went on to write a Master's thesis in which he showed how Boolean algebra could optimize the design of systems of electromechanical relays then used in telephone routing switches. He also proved that circuits with relays could solve Boolean algebra problems. Employing the properties of electrical switches to do logic is the basic concept that underlies all modern electronic digital computers. Thus Boole, via Shannon, provided the theoretical grounding for the Digital Age.

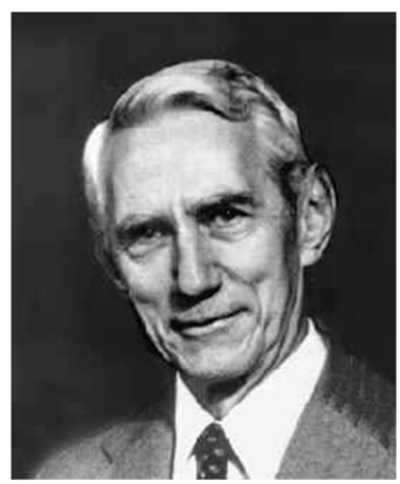


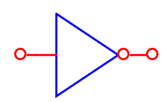
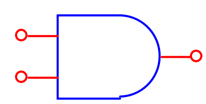
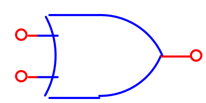
Figure 10.1: Claude Shannon
1916-2001

10.3 Boolean Operations

Boole denoted the universe of all things by unity. If all things are denoted by unity then the inverse of all things is nothing, denoted by zero. There are three basic operators in Boolean algebra, NOT, AND, and OR. We can define our universe (or the set of all things) however we want. For example we can define our universe as all students. If we then define X = all EE102 students, then NOT(X) = students not in EE102. If X = EE102 students and Y = Math251 students, then X AND Y would be all EE102 students who are also taking Math251. X OR Y would be all students taking either EE102 or Math251. For instance if Joe is taking EE102 then X=1 or true, Joe is an EE102 student. If Joe is not taking Math251, then Y=0 or false. X AND Y would then equal 0 or false. Joe is not both an EE102 student AND a Math251 student. X OR Y would equal 1 or true. Joe is taking either EE102 OR Math251.

We denote those electrical circuits that perform the NOT, AND, and OR Boolean operations with the symbols shown in Table 10.1. The triangle symbol is just a buffer, i.e. what goes in comes out. The bubble on the output indicates that what comes out is inverted from what goes in (i.e. NOT). The other two symbols represent the AND gate (middle) and OR gate (bottom). If either of those symbols has a bubble on its output then those gates would become a NAND (or NOT AND) and a NOR (or NOT OR) gates respectively.

Table 10.1: Boolean Operations, definition and gate symbol

Boolean Operations		Gate Symbol
NOT	$NOT(X) = \overline{X} \Rightarrow \overline{0} = 1 \quad \overline{1} = 0$	
AND	$AND(X,Y) = X \cdot Y \Rightarrow \begin{matrix} 0 \cdot 0 = 0 & 0 \cdot 1 = 0 \\ 1 \cdot 0 = 0 & 1 \cdot 1 = 1 \end{matrix}$	
OR	$OR(X,Y) = X + Y \Rightarrow \begin{matrix} 0 + 0 = 0 & 0 + 1 = 1 \\ 1 + 0 = 1 & 1 + 1 = 1 \end{matrix}$	

10.4 Boolean Algebra

A Boolean function $F(X_1, X_2, \ldots X_N)$ is a function of N Boolean variables or inputs whose output is also Boolean. It consists of an algebraic expression formed with the N variables, the constants 0 and 1, and the logic operation symbols: NOT, AND, and OR. For example we can write our Boolean equations as

$$F_1 = X \cdot \overline{Y} + \overline{X} \cdot Z \qquad \text{(x.x)}$$

In this expression there are three Boolean variables, X, Y, and Z. The overbar represents the NOT operation, the plus the OR operation, and the dot the AND operation. There are eight possible combinations of three input variables. The truth table shown in Table 10.2 allows us to easily see what the output will be for each combination of those three variables. For instance if X, Y, and Z were each zero, then the expression would be

$$
\begin{aligned}
F_1 &= X \cdot \overline{Y} + \overline{X} \cdot Z \\
&= 0 \cdot \overline{0} + \overline{0} \cdot 0 \\
&= 0 \cdot 1 + 1 \cdot 0 \\
&= 0 + 0 \\
&= 0
\end{aligned}
\qquad \text{(x.x)}
$$

The order of precedence is to perform the NOT first, then the AND, then finally the OR. So NOT(0)=1, $0 \cdot 1=0$, $0 + 0 = 0$, therefore $F_1 = 0$.

Table 10.2: Truth table for function defined by (x.x)

X	**Y**	**Z**	**F_1**
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	0

I can determine when $F_1 = 1$ by individually evaluating each term of the OR operation. When either term equals 1 then $F_1 = 1$. So if X and not Y = 1 then F1 = 1. This implies that $F_1 = 1$ when X = 1 and Y = 0. Also for Not X and Z = 1 when X = 0 and Z = 1. $F_1 = 0$ for all other combinations. An alternative manner to represent the NOT operation is to use an apostrophe. For example F_1 could be written as:

$$F_1 = X \cdot Y' + X' \cdot Z \qquad \text{(x.x)}$$

As you learned in high school, algebra has many identities that define algebraic operations. Boolean Algebra has a similar set of identities (see Table 10.3). Some of them are the same as for regular algebra, such as Associative, Commutative, and Distributive laws. Some are unique to Boolean Algebra such as DeMorgan, Simplification, and Absorption. Using these identities we can prove that two functions are identical or simplify a complex function.

Table 10.3: Boolean Algebra Identities

$X + 0 = X$	$X \cdot 1 = X$	Identity
$X + 1 = 1$	$X \cdot 0 = 0$	Null
$X + X = X$	$X \cdot X = X$	Idempotence
$X + \overline{X} = 1$	$X \cdot \overline{X} = 0$	Complementary
$\overline{\overline{X}} = X$		Involution
$X + Y = Y + X$	$X \cdot Y = Y \cdot X$	Commutative

$X+(Y+Z)=(X+Y)+Z$	$X\cdot(YZ)=(XY)\cdot Z$	Associative
$X\cdot(Y+Z)=XY+XZ$		Distributive
$\overline{X+Y}=\overline{X}\cdot\overline{Y}$	$\overline{XY}=\overline{X}+\overline{Y}$	DeMorgan
$X+\overline{X}Y=X+Y$	$X\cdot(\overline{X}+Y)=XY$	Simplification
$X+XY=X$	$X\cdot(X+Y)=X$	Absorption

For instance, how do we prove that the following equality is correct?

$$X\cdot\overline{Y}\cdot Z+\overline{(X\cdot Z)}=\overline{Y}+\overline{(X\cdot Z)} \quad \text{(x.x)}$$

First, let's define a new variable $W=XZ$. Substituting into equation (x.x) and using the Simplification Identity gives

$$\begin{aligned} W\cdot\overline{Y}+\overline{W}&=\overline{Y}+\overline{W} \\ &=\overline{Y}+\overline{(XZ)} \end{aligned} \quad \text{(x.x)}$$

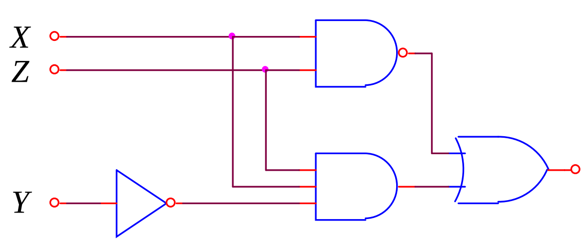
showing that equation (x.x) is correct.

We can simplify or optimize equations in the same way. For example:

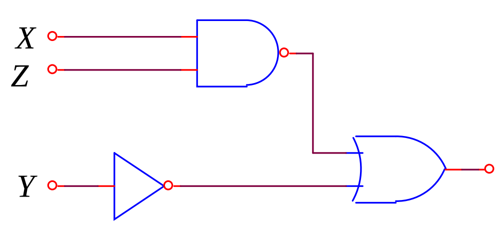
$$\begin{aligned} XY+\overline{X}+Y\overline{Z}&= \\ &=Y+\overline{X}+Y\overline{Z} && \text{Simplification} \\ &=Y+\overline{X} && \text{Absorption} \end{aligned} \quad \text{(x.x)}$$

10.5 Logic Gates

There is a direct relationship between the Boolean expression that I wish to implement and the connections of logic gates that I use. For example, consider equation (x.x).



(a)



(b)

Figure 10.3: Direct logic gate implementation (a) left side and (b) right side of equation (x.x).

When working with "ideal" logic gates there are certain points to remember.

- You are allowed to have any number of inputs for any gate.
  For real discrete gates that is not really true although you can purchase two-input easily and can even find some 8 input gates. In programmable logic there are many more inputs but not really infinite.
- Any input to a gate is either an input to the circuit or an output of another gate.

This just means that all inputs must be specified in some way otherwise the output is unknown.

- Two outputs cannot be connected together to the same input of another gate.
  Connecting two outputs to a single input could potentially create a "bus contention" where one output is trying to drive the input high and the other output is trying to drive the input low. The outputs of "open collector" gates, however, may be connected creating what is called a "WIRED OR" or a "WIRED AND" by providing an appropriate pull down or pull up resistor.
- An output can be the input to as many gates as desired. This is true only for ideal gates. Real gates all have "fan out" limitations. However, typical fan out limits usually exceed the desired number of gates that most people wish to drive.
- Values, i.e. high or low, 1 or 0, VCC or GND, will be assigned to inputs and never to outputs. You should never tie an output directly to VCC or GND.

10.6 Truth Tables - Minterms

As we have previously seen the truth table fully describes a Boolean function in terms of its input variables. Each row of the truth table can be defined as a product of the input variables or terms that causes the function to be 1 or true. That product is called a minterm.

Table 10.4: Truth table of an arbitrary function with three input variables showing the associated minterms.

Row	***X***	***Y***	***Z***	***F*₂**	**Minterm**
0	0	0	0	0	$m_0 = \overline{X} \cdot \overline{Y} \cdot \overline{Z}$
1	0	0	1	0	$m_1 = \overline{X} \cdot \overline{Y} \cdot Z$
2	0	1	0	1	$m_2 = \overline{X} \cdot Y \cdot \overline{Z}$
3	0	1	1	1	$m_3 = \overline{X} \cdot Y \cdot Z$
4	1	0	0	0	$m_4 = X \cdot \overline{Y} \cdot \overline{Z}$
5	1	0	1	1	$m_5 = X \cdot \overline{Y} \cdot Z$
6	1	1	0	0	$m_6 = X \cdot Y \cdot \overline{Z}$
7	1	1	1	0	$m_7 = X \cdot Y \cdot Z$

For example, consider Table 10.4 row 2: X = 0, Y = 1, and Z = 0. In order for a function, F_2, consisting of a product of these terms to be 1 given their input, the product would need to be: $\overline{X} \cdot Y \cdot \overline{Z}$. Each row in the truth table can be represented by its associated minterm (see Table 10.4).

The function which represents any given truth table is just the sum of the minterms associated with the 1's in the truth table. The function for the truth table shown in Table 10.4 has three 1's (row's 2, 3, and 5). The equation which represents this function is the sum of the minterms m_2, m_3, and m_5.

$$F_2 = \overline{X}Y\overline{Z} + \overline{X}YZ + X\overline{Y}Z \qquad \text{(x.x)}$$

A shorthand notation for this summation is

$$F_{XYZ} = \sum(2,3,5) \qquad \text{(x.x)}$$

Now that the Boolean equation for the function has been determined we can develop a two-level AND-OR circuit that will implement this function (see Figure 10.4). Note that the input to this two-level circuit, called the sum-of-products or SOP, is every possible combination of X, Y, and Z and their inverses: $\overline{X}$, $\overline{Y}$, and $\overline{Z}$.

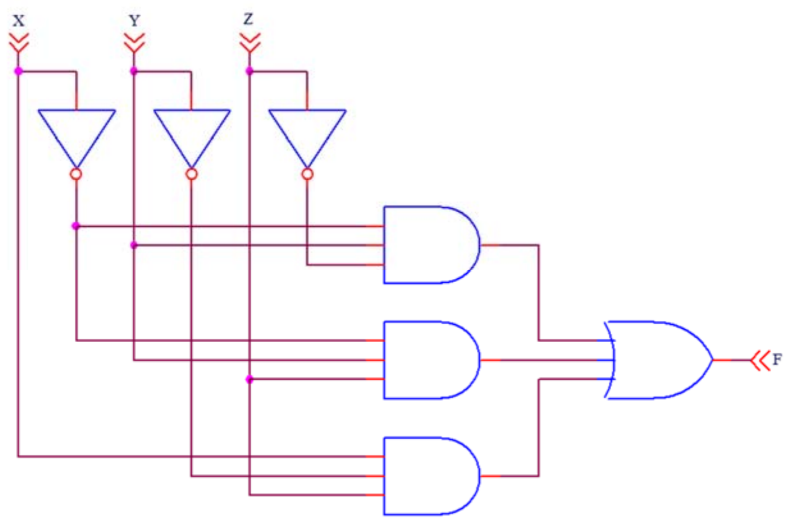


Figure 10.4: Two-level sum-of-products implementation of the function given in equation (x.x)

All truth tables can be represented by a two-level sum-of-products circuit. Although creating circuits using the two-level sum-of-products method is fairly straight forward, it does not necessarily create the most optimal circuit in terms of minimizing the number of gates and inputs necessary to create the same function.

10.7 Optimizing Circuits – Karnaugh Maps

Consider the Boolean equation (x.x). We would like to optimize this equation to minimize the number of gates and inputs to those gates in the resultant circuit. One approach is to apply Boolean Algebra identities to minimize the equation. For instance, the first two terms can be reduced as follows.

$$\begin{aligned} \overline{X}Y\overline{Z} + \overline{X}YZ &= \\ &= \overline{X}Y \cdot (\overline{Z} + Z) && \text{Distributive} \\ &= \overline{X}Y \cdot 1 && \text{Complementary} \\ &= \overline{X}Y && \text{Identity} \end{aligned} \qquad \text{(x.x)}$$

Therefore, the first two products of three variables is reduced to a single product of two variables.

A Karnaugh Map is a graphical tool to perform similar circuit minimization. Basically a Karnaugh Map is just a truth table rearranged so that for each adjacent cell only one input variable changes state.

For example, the rows of the truth table (Table 10.4) are arranged in the Karnaugh map in the following fashion (see Figure 10.5) where the numbers represent the possible minterms of the function. Notice that the rows or minterms of the truth table are mixed up in the Karnaugh map. So minterm 1 is next to minterm 3. The Karnaugh Map is specifically arranged so that if you transition from one cell to the next only one input variable changes its state. For example moving from m_1 to m_3, variable Y changes for a 0 to a 1 while X remains 0 and Z remains 1.

X \ YZ	00	01	11	10
0	m_0	m_1	m_3	m_2
1	m_4	m_5	m_7	m_6

Figure 10.5: Karnaugh map showing the minterms of a truth table with variablex X, Y, and Z.

Consider Table 10.2. The sum of products representation of this truth table is

$$F_1 = \sum_{XYZ}(1,3,4,5) = \overline{X} \cdot \overline{Y} \cdot Z + \overline{X} \cdot Y \cdot Z + X \cdot \overline{Y} \cdot \overline{Z} + X \cdot \overline{Y} \cdot Z \qquad \text{(x.x)}$$

The Karnaugh map for Table 10.2 is shown in Figure 10.6.

X \ YZ	00	01	11	10
0		1	1	
1	1	1		

Figure 10.6: Karnaugh map for Table 10.2

Remember that the minterms in the equation represent where the 1's are in the truth table, so we can put 1's in the cells of the Karnaugh map that represent these minterms.

On a Karnaugh Map we indicate combining terms by drawing a ring around the terms that, when combined, yield a simpler expression. In this example we can draw a ring around m_1 and m_3. Comparing these two squares, we see that the variable that changes from 0 in m_1 to 1 in m_3 is variable Y; hence, when the two minterms are combined, this variable is eliminated. The reduced term is $\overline{X} \cdot Z$. Combining m_4 and m_5 yields $X \cdot \overline{Y}$. The minimized function is $X \cdot \overline{Y} + \overline{X} \cdot Z$, which is what we started with in equation (x.x).

The general guidelines for simplifying functions using Karnaugh Maps are as follows:

- When combining terms (squares) on a Karnaugh map, always group the squares in powers of 2, that is 2, 4, 8, etc. Grouping 2 squares eliminates one variable, grouping 4 squares eliminates 2 variables, etc. In general grouping 2^N squares eliminates N variables.

- Group as many squares together as possible; the larger the group is the fewer the number of variables in the resulting product term.
- Make as few groups as possible to cover all the squares or minterms of the function. A minterm is covered if it is included in at least one group. The fewer the groups, the fewer the number of product terms in the minimized function.
- Each minterm may be used as many times as it is needed; however, it must be used at least once. As soon as all the minterms are used once, stop.

Consider the following Karnaugh Map. Using the above guidelines we can group the minterms as follows. Note that the opposite edges of the map are actually coincident. Therefore m_0, m_1, m_8, and m_9 may be grouped together. The largest grouping which includes m_{15} combines it with m_7. m_3 may be grouped with m_7 or with m_1. Either grouping is valid, but only one is needed.

YZ WX	00	01	11	10
00	1	1	1	
01			1	
11			1	
10	1	1		

Figure 10.7: Example Karnaugh map

For each grouping we can write down the resultant simplified term by identifying those variables that do not change. For example, the grouping of m_0, m_1, m_8, and m_9, W is both 0 and 1, X is always 0, Y is always 0, and Z is both 0 and 1. Therefore only X and Y are retained in the resultant combination ($\overline{X} \cdot \overline{Y}$). For the grouping of m_7 and m_{15}, only W is both 0 and 1, X, Y, and Z are always 1, so the combined terms is $X \cdot Y \cdot Z$. For the grouping of m1 and m3, Y is both 0 and 1, W and X are 0 and Z is 1, so the combined terms is $\overline{W} \cdot \overline{X} \cdot Z$.

Think About It

What is the minimized function for the following Karaugh Map.

YZ WX	00	01	11	10
00	1		1	
01	1	1		
11	1	1		
10	1			1