
PHYSICS LAB TECHNIQUES

PHYS 497-597

APRIL 5, 2022

ELIJAH THOMPSON,
PHYSICS AND MATH HONORS

Solo Pursuit of Learning



Contents

Part I

Laboratories

Chapter 1

Pendulum

1.1.0 Background

A simple pendulum is a mass attached to a string. The time needed for one full oscillation (period, T) is related to the length of the pendulum L and acceleration due to gravity g . The relationship is given by the following equation:

$$T = 2\pi \sqrt{\frac{L}{g}} \quad (1.1.1)$$

The pendulum can be used to determine the acceleration due to gravity by measuring its length and period of oscillation.

In each part of the experiment you will measure the length of the pendulum and its period (with uncertainties) in order to calculate the acceleration due to gravity and determine its uncertainty. The four values of g you will find will be compared to the expected value within uncertainties. The accuracy of each part of the experiment will be determined.

1.2.0 Experimental Procedure

1.2.1 Preliminary Measurements

First we determine the human reaction time of the measurers.

Step 1: Start the stopwatch and attempt to stop it exactly at $t = 5.00\text{ s}$. Record the value of the difference between the reading and 5.00 s mark in Excel. Repeat three times.

Step 2: Calculate the average reaction time of your lab section and the standard deviation associated with the distribution. This data will help you to properly determine the uncertainty in the time of the pendulum swing.

Next we measure the length of the pendulum. First appropriately define the length of your pendulum (where do you start and stop your measurement).

1.2.2 Main Data Collection

Set up the pendulum as shown in the Figure 1.1.

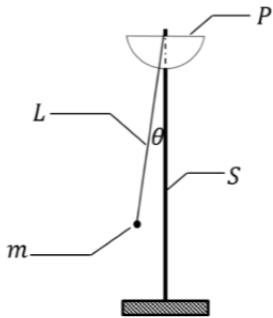


Figure 1.1: Experimental Setup.

Next perform singular measurements of oscillation times:

Step 3: Select a string of length of approximately 1.0 m. Attach the selected mass to the string. Measure the length L of the pendulum and record it in Excel. Estimate an uncertainty in one measurement taking into account everything discussed in preliminary measurements.

Step 4: Measure the time T for one oscillation of the pendulum. Make sure to use a small angle (less than 10 degrees) for all oscillations, as we use the small angle approximation. Record that value and determine its uncertainty $u(T)$ in Excel (use the information about the reaction time from the preliminary measurements)

Step 5: Using the oscillation time T and the length of the pendulum L , rearrange equation 1.1.1 to calculate the acceleration due to gravity g_1 . Use uncertainty propagation to calculate uncertainty $u(g_1)$.

Multiple measurements of an oscillation time:

Step 6: Repeat the measurement of a single oscillation ten times. Record the values in your Excel. Estimate the uncertainty of each time measurement.

Step 7: Determine the average value of the period of the pendulum \bar{T} and its uncertainty.

In this part of the lab we are deriving a Type A uncertainty for the period. Does this encapsulate all of your uncertainty about the measurement?

Step 8: Using the Average period of the oscillation \bar{T} and the length of the pendulum L determined in the previous section, calculate the acceleration due to gravity g_2 and its uncertainty $u(g_2)$.

Singular measurement of time of multiple oscillations:

Step 9: Measure the time t of ten oscillations of the pendulum and estimate the uncertainty in your time measurement. Record your results in Excel.

Step 10: Using the time t of ten oscillations, calculate the period of one oscillation T with the uncertainty $u(T)$. Use the uncertainty propagation to calculate $u(T)$.

Step 11: Using the oscillation time T and the length of the pendulum L from the first section, determine the acceleration due to gravity g_3 with uncertainty $u(g_3)$.

Measurements of multiple oscillations for various pendulum lengths:

Step 12: Discuss the six different values of length you are going to use and why. You can all use the same lengths or different ones. Discuss the advantages of each approach and justify your decision in the report.

Step 13: Set up a pendulum of a given length L . Record the length and its uncertainty.

Step 14: For each length, measure the time of ten full oscillations of the pendulum. Perform this experiment only once for each length. Record the values in Excel. Estimate the uncertainty of each time measurement.

When plotting our data we want it linearized in the form $y = mx + b$, so if we want period versus time we can plot $y = T$ and $\sqrt{L} = x$ so that $m = 2\pi/\sqrt{g}$.

Chapter 2

Pendulum II

2.1.0 Background

In the previous experiment the angular position of the pendulum was modeled using simple harmonic motion and it was assumed to depend on time according to the equation

$$\theta(t) = \theta_0 \cos(\omega t + \theta_i) \quad (2.1.1)$$

where θ_0 is the maximum angular displacement of the pendulum from its equilibrium position, θ_i is the initial displacement angle, and ω is the angular frequency of the oscillation. Equation 2.1.1 is a solution of the ordinary differential equation

$$\frac{d^2\theta}{dt^2} = -\frac{mgd}{I}\theta \quad (2.1.2)$$

where I is the moment of inertia of a pendulum about the pivot point, m is the mass of the pendulum, g is the acceleration due to gravity, and d is the distance between the axis of rotation and the pendulum's center of mass. For the simple pendulum consisting of a string of negligible mass and a small bob of mass m the moment of inertia $I = mL^2$ and $d \approx L$. Equation 2.1.2 simplifies to

$$\frac{d^2\theta}{dt^2} = -\frac{g}{L}\theta = -\omega^2\theta \quad (2.1.3)$$

From that we can conclude that the period of the pendulum undergoing simple harmonic motion can be written as:

$$T = \frac{2\pi}{\omega} = 2\pi \sqrt{\frac{L}{g}} \quad (2.1.4)$$

All of these equations are results of the approximation $\sin \theta \approx \theta$, which is valid for θ measured in radians when $\theta < 10^\circ \approx \frac{\pi}{16}$.

This experiment we will consider the measurement of the period of a simple pendulum that is not constrained by the small angle approximation. We will investigate the limitations of the model, the experimental set up, and data collection procedures.

2.2.0 Experimental Procedure

2.2.1 Preliminary Measurements

Step 1: Using your experience from last week, design an experiment that will allow you to test the validity of the small angle approximation, i.e., a set up and data collection procedure that will allow measurement of the period with high precision and accuracy so it would be possible to detect whether Equation 2.1.4 holds for larger amplitudes:

1. Write down a plan for a high-precision measurement of the period of the pendulum at different amplitudes and for different pendulum sizes.
2. Provide a clear description on the method that will be used for determining the uncertainty in the measurement.

Step 2: Set up the designed experiment using the equipment provided.

Step 3: Test your experimental setup and the method of obtaining the period of the pendulum by performing the experiment for a small angle first.

Step 4: Check the quality of your data. Compare the value of the period to the values obtained previously. Redesign the experimental set up if issues arise.

2.2.2 Main Data Collection

First we perform a period dependence investigation:

Step 5: After committing to the experimental set up, collect data that shows the dependence of the period of the pendulum on the initial amplitude of the oscillation ($T(\theta_i)$). Obtain as large of a range of θ_i as you can possibly get.

Step 6: Using Excel, plot $T(\theta_i)$ using your data and a theoretical plot $T(\theta_i)$ using Equation 2.1.4 for your pendulum.

Step 7: Discuss similarities and discrepancies between the experimental and theoretical curves.

Now look at your initial experiment as a scientist measuring an unfamiliar phenomenon. This will require you to focus on one or more aspects of the deviation from the small angle model.

Step 8: Investigate one of the aspects of the model listed below:

1. The issue of $\theta_{critical}$: How precisely can you determine where small angle approximations break down? Consider the precision of the measurement, statistical distributions, and any other factors that could affect your measurement.

2. The issue of θ_{max} : As a physical system, the mass on a string will be affected by factors disregarded in the model. What are the physical limitations of the measurement? What is the maximum range of angles for which $T(\theta_i)$ can be determined?
3. The issue of theoretical T . As Equation 2.1.4 is no longer valid for large θ_i , it is unlikely the relationship between T and L will be the same. What would that relationship look like? What other factors would the period of the pendulum depend on?

Step 9: Formulate a hypothesis for any branch. Like all hypotheses, justification and logical processes that lead to the hypothesis should be included.

Step 10: Plan your experiment. Clearly indicate all the measurements you will take.

Step 11: Set up your experiment and collect data.

2.3.0 Data Analysis

Step 12: Plot your data to show the period dependence on the amplitude

Step 13: Discuss similarities and discrepancies between the experimental and theoretical curves

Step 14: Plot any additional data collected.

Step 15: Search for existing theoretical work and models deriving an equation for the period of the pendulum without the small angle approximation. Properly reference sources in *Canadian Journal of Physics* style.

Step 16: Compare the theoretical results and trends with those shown by your data.

Chapter 3

AC Measurements and Sources



Figure 3.1: A photo of the experimental setup.

3.1.0 Background

Almost all devices operate and communicate using time-varying electrical signals and therefore, the ability to measure the properties of these signals (amplitude, time variation, etcetera) is critical. The DC signals (signals that do not vary in time) are generated with a power supply while AC signals (signals that are time varying) are generated with a function generator in this lab.

3.1.1 DC Voltage Measurements

A multimeter is a good choice for measuring a steady voltage, such as that across the terminals of a power supply, but an oscilloscope can also be used, provided that great accuracy is not required. It is instructive to measure the terminal voltage of a power supply with both a multimeter and an oscilloscope and compare results as in Figure 3.2.

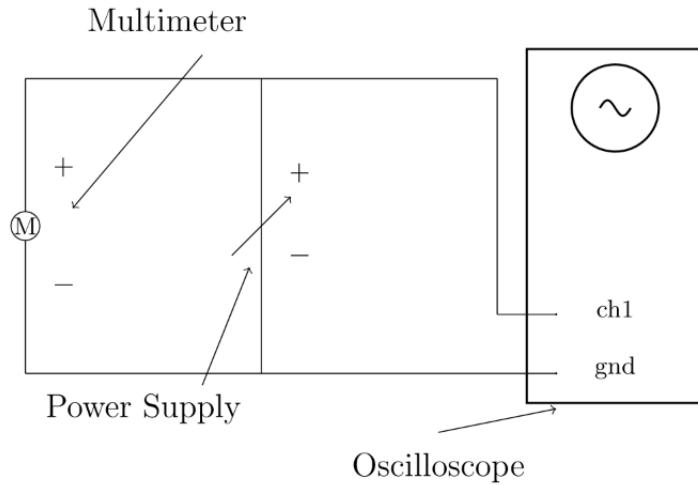


Figure 3.2: Measuring DC power supply with oscilloscope and multimeter.

Both the reading on the multimeter and the displacement of the trace on the oscilloscope (difference between the position of the trace at *ground* and when the source is connected) represent the voltage provided by the power supply. Note that the oscilloscope trace only moves when the input coupling control is set to DC. In AC coupling mode the DC component of the input signal is blocked and the trace will remain stationary.

Note 3.1.1. The oscilloscope only measures voltage, so when current is being measured it should be removed from the circuit.

3.1.2 AC Voltage Measurement

An oscilloscope is most useful for measuring time varying signals. Replace the power supply in the circuit with a sine function generator as seen in Figure 3.3

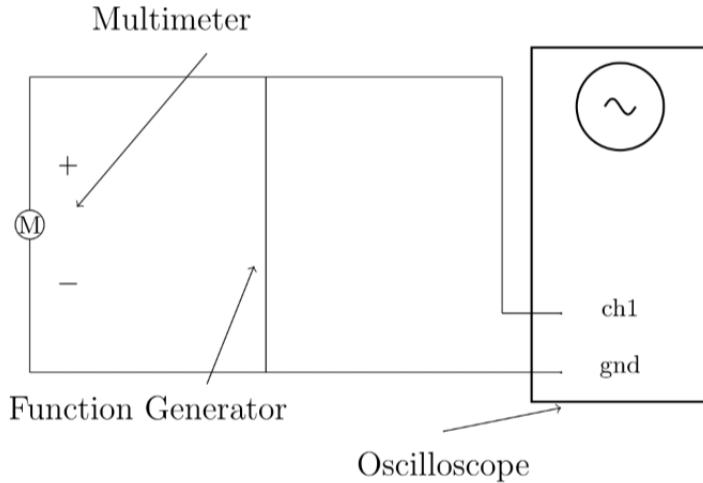


Figure 3.3: Measuring AC power supply with oscilloscope and multimeter.

The multimeter reading can now be compared to the oscilloscope display. The trace on the display is a sine wave $V(t)$ whose equation is given by

$$V(t) = V_p \sin(\omega t) = V_p \sin(2\pi ft) = V_p \sin\left(\frac{2\pi}{T}t\right) \quad (3.1.1)$$

where ω is the angular frequency of the wave in radians/second, f is the frequency of the wave in Hertz, T is the period of the wave in seconds, and t is the time in seconds.

There are four ways of describing the magnitude of the AC waveform:

- a) The **peak value** V_p is the maximum voltage of the waveform.
- b) The peak to peak value V_{pp} is the full voltage difference between maximum and minimum values of the voltage. For the sinusoidal wave given by Equation 3.1.1 this equals $2V_p$
- c) The average value V_{av} over half a period. This value is found from the equation

$$V_{avg} = \frac{2}{T} \int_0^{T/2} V(t) dt = \frac{2}{T} \int_0^{T/2} V_p \sin(\omega t) dt = \frac{2V_p}{\pi} \quad (3.1.2)$$

- d) The effective or **root mean square** (RMS) value, V_{rms} , is defined as the equivalent DC voltage that will supply the same power, P , to a resistance, R , as the original waveform over a full period. From the definition of power:

$$P = \frac{1}{T} \int_0^T \frac{V^2(t)}{R} dt = \frac{1}{T} \int_0^T \frac{V_p^2}{R} \sin^2(\omega t) dt = \frac{V_p^2}{2R} \quad (3.1.3)$$

By the above definition, the RMS voltage is that DC voltage which would supply the same power to the resistor. Therefore,

$$P = \frac{V_{rms}^2}{R} = \frac{V_p^2}{2R} \quad (3.1.4)$$

and thus

$$eq : ACDC5V_{rms} = \sqrt{RP} = \sqrt{\frac{1}{T} \int_0^T V^2(t) dt} = \sqrt{\frac{1}{T} \int_0^T V_p^2 \sin^2(\omega t) dt} = \frac{V_p}{\sqrt{2}} \quad (3.1.5)$$

These different **magnitudes** can be seen in the sine wave shown in Figure 3.4.

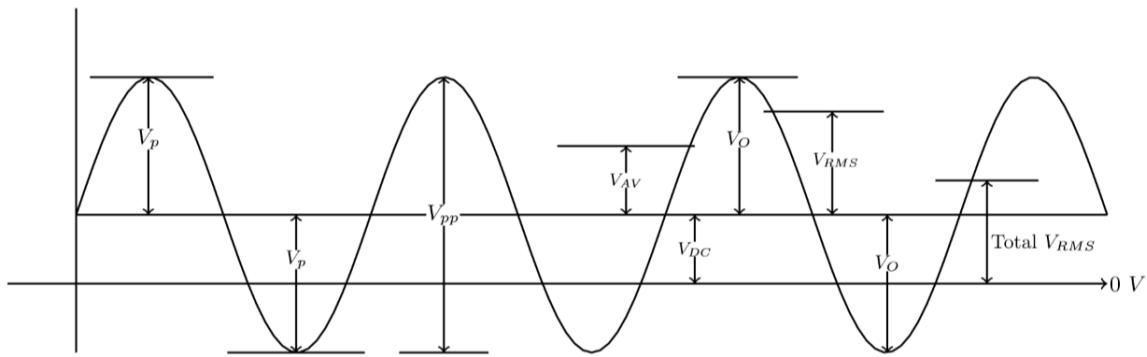


Figure 3.4: The four ways of defining the amplitude of a sine wave.

3.1.3 AC and DC Measurements

The offset control on the function generator can be used to add a DC component to the output waveform. This can be seen in Figure 3.4 where the wave is offset vertically. The sine wave can be observed to move up and down on the oscilloscope when the offset control is adjusted. As with the power supply, one can see that the trace will only move when the input controls are set to DC coupling. With AC coupling the DC level is removed and only the sine wave appears on the screen. Similarly, when the multimeter is set to read AC volts or amps, only the AC component is measured. When the multimeter function is DC volts or amps, only the DC level is measured.

The equation of a sinusoidal wave together with a DC level is given by

$$V(t) = V_p \sin(\omega t) + V_{DC} \quad (3.1.6)$$

Here the frequency is the same as in Equation 3.1.1 but the values of $V(t)$ are different. The sinusoidal wave with extra DC component would supply more power to a resistor compared to the wave with $V_{DC} = 0$ V. For the case of a waveform riding on top of a DC level, the RMS amplitude is given by the formula

$$\text{Total } V_{rms_{DC}} = \sqrt{V_{rms}^2 + V_{DC}^2} \quad (3.1.7)$$

which for the wave described by Equation 3.1.6 reduces to

$$\text{Total } V_{rms_{DC}} = \sqrt{\frac{V_p^2}{2} + V_{DC}^2} \quad (3.1.8)$$

3.1.4 The Function Generator

The function generator is an instrument that generates time varying electrical signals. The front panel controls set the shape, size, rate, and other properties of the output waveform. The BK 3011B can generate sine waves, triangle waves, and rectangular waves in a frequency range from 0.2 Hz to 2.0 MHz.

Before connectin the function generator into a circuit, it is essential to make sure that it has been properly configured for the appropriate output signal. Standard practice is to use an oscilloscope to view the output of the function generator directly. Once the waveform is correctly displayed on the screen, the function generator is powered off and connected to the circuit. Do not connect an operating function generator to a circuit as it is being wired together. When the circuit wiring is completed, the function generator is powered up, along with any other power supplies and sources that the circuit needs. The function generator cannot withstand a short circuit across the output terminals. Always check the circuit you are driving with the function generator to make sure that the output is not being shorted out.

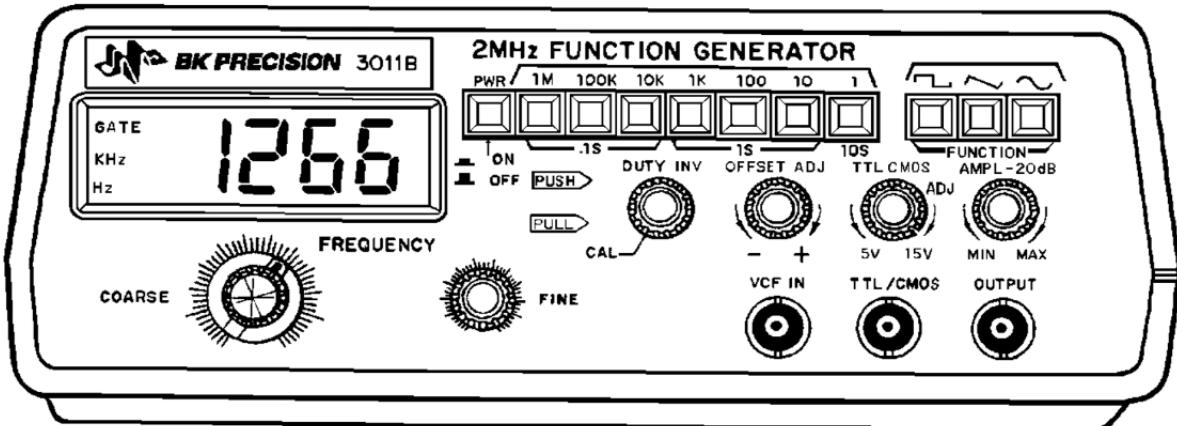


Figure 3.5: BK 3011B function generator control panel.

Figure 3.5 shows the operating controls of the function generator. When setting up this instrument, the first task is to choose the appropriate shape (sine, square, or triangle) using the three **FUNCTION** buttons. The waveform appears at the connector labeled **OUTPUT**. The other two output jacks are not commonly used. Secondly, adjust the amplitude of the signal by adjusting the knob labeled **AMPL** to about one quarter of the way between **min** and **max**. The buttons labeled **DUTY**, **OFFSET**, **TTL**, and **AMPL** must all be pushed in. Pulling them out changes the range or adds other functions. The left knob of these four, **DUTY**, changes the symmetry of the waveform. Make sure this knob is rotated fully counterclockwise to the **CAL** position. If it is not in the proper position, some output waveforms will appear distorted.

Lastly, it remains to set how often the waveform repeats itself. The push buttons labeled 1, 10, 100, 1K, 10K, 100K, and 1M set the **range** of the repetition rate. Inside any given range the repeat rate is set by the **COARSE** and **FINE** controls. For sine waves these controls adjust the **frequency** of the wave. For triangle wave and rectangle waves, the frequency is not an accurate description and the term **repetition rate** is used instead.

3.1.5 The Oscilloscope

An **oscilloscope** is an instrument which usually displays a graph (trace) of the voltage of a signal plotted against time. Time is plotted horizontally, in the *x*-direction and voltage is plotted vertically, in the *y*-direction. The graph changes as the voltage changes, so both the waveform and any changes in the waveform can be monitored. Both the voltage scale and the time scale can be adjusted over several orders of magnitude. Measurements of voltages and times can be made by comparing the trace with the transparent scale mounted in front of the screen on which the trace appears.

The trace is formed by the motion of a bright spot which appears where a narrow electron beam hits the phosphor on the screen of a cathode ray tube (CRT). The electron beam can be deflected in the *x* and *y* directions by the electric field caused by amplified signals applied to deflector plates inside the tube. These amplified signals are generated inside the oscilloscope in response

to controls the operator sets as well as the actual waveform to be measured.

Figure 3.6 shows the controls of a typical oscilloscope.

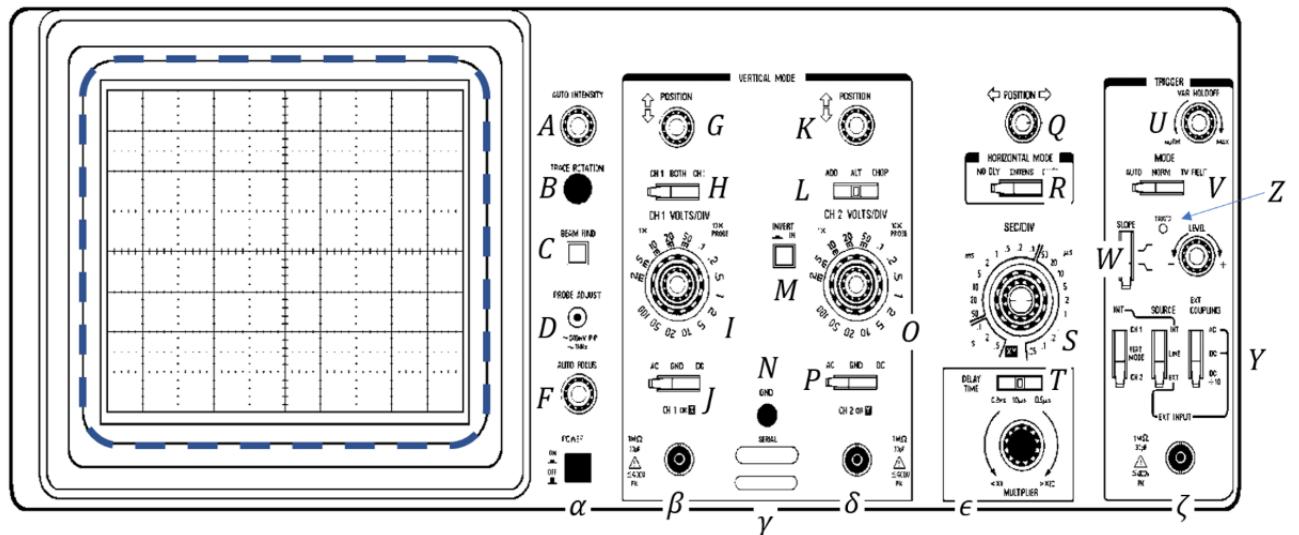


Figure 3.6: Oscilloscope display panel. A - auto intensity, B - trace rotation, C - beam find, D - probe adjustments, F - auto focus, G and K - Position controls for CHANNEL 1 and CHANNEL 2, respectively, H and L - vertical mode switches (CH1, BOTH, CH2, ADD, ALT, CHOP), I and O - 1X and 10X probes for VOLTS/DIV range for CHANNEL 1 and CHANNEL 2, respectively, J and P input coupling (AC, GND (Ground), DC), M - invert, N - ground connector, Q - position control for horizontal display, R - horizontal mode switch, S - SEC/DIV switches, T - delay time range selector, U - variable holdoff control, V - more switch (auto, norm, and TV field), W - slope switch, X - level control, Y - external coupling (L), source of trigger signal (M) and internal source trigger source (R), Z - external input connector, α - power switch, β and δ - serial and mod slots, γ - serial and model number slots, ϵ - delay time multiplier ζ .

We now describe the controls seen in Figure 3.6

1. Basic display controls:

- **Focus**: This controls the size of the spot on the screen, and thus sharpness of the trace which forms the graph. It should be adjusted to give a clear fine image.
- **Intensity**: This controls the brightness of the spot. If the spot is too dim the trace is too faint to see. A trace which is too bright will eventually leave a permanent mark on the screen, and in any case tends to smear the trace and hide details. It is good practice to use the lowest intensity that will show all necessary detail in the trace without straining the eyes.

2. Position controls:

These control the horizontal and vertical position of the trace. Usually only small changes are necessary in the horizontal position of the trace. Quite large adjustments may be needed for the vertical position. If no trace is visible on the screen, reduce the vertical

gain or ground the signal using input coupling controls. Then make sure that the intensity is not too low, and set the *x* and *Y* position controls to the middle of their range. It may be that the time base is not triggering. Some oscilloscopes have a “beam find” button which can be pushed to locate the trace if it off the screen.

3. Vertical Gain:

The vertical scale of the display is set by the vertical gain control (***SYMBOL***), usually a multi-position rotary switch. The scale is usually given in V/cm or mV/cm . A dual trace oscilloscope has two independent sets of controls, one for each channel. If a continuous control is provided, usually a small knob in the center of the rotary switch, make sure it is turned completely to the ***CAL*** position before making any measurements. Record the voltages for each channel independently as they may be set to different scales.

4. Coupling Controls:

The input signal can be connected (“coupled”) to the amplifier for the vertical channel(s) either directly (DC coupling) or through a capacitor (AC coupling). A third possibility is to ground the input of the amplifier, which provides a quick way of finding the zero signal position of the trace. AC coupling blocks any DC component of the signal. This is useful when a varying signal is superimposed on a steady voltage, and one wants to zoom in to examine the varying part in detail. If the steady part of the signal needs to be examined then DC coupling is required.

5. Time Base:

The time base or sweep moves the spot horizontally to the right at a controlled speed, then returns it rapidly to the left of the screen to start again. The sweep or time-base sets the horizontal scale of the display, and is typically controlled by a multi-position rotary switch. The scale is given in s/cm , ms/cm , $\mu s/cm$, and ns/cm . There may be a knob at the center of the rotary switch which gives a continuous change of scale, but which clicks to a preset “calibrated” setting when turned to the far left or right. This control must be in the “calibrated position” before making any measurements of time intervals. It is good practice to check the appearance of the signal at a wide range of time-base settings before starting to make measurements. Some features are best seen at slow speeds while others may show up better at high speeds.

6. Triggering:

This controls the time at which the sweep starts. For the display to be steady the sweep should start at the same point on a repetitive signal every time the sweep begins. In this way successive traces will be in the same position:

- ***Trigger Source:*** The choices are usually INT (an internally generated signal being supplied to the horizontal amplifier), EXT (some external signal source), and LINE (the 60 Hz line frequency). For most purposes INT is the correct choice. Some signal generators have a SYNC output signal which can be used to trigger the oscilloscope using EXT. The LINE trigger source is used for observing signals that are related to the AC power line. If there is a trigger mode control usually the AUTO setting is the correct one. With dual trace oscilloscopes make sure the trigger source is the appropriate channel. A channel with no signal will give poor triggering.

- **Trigger Level:** This sets the signal voltage value at which the sweep starts. If an AUTO mode is available, try it. If this gives a good trace use that setting. Otherwise, experiment with the LEVEL control. If the LEVEL setting is too low, the sweep will trigger at several different points in the waveform, and the successive traces will not overlap properly, giving an unstable messy display. If the level is too high, the sweep may not trigger at all, and there will be no trace, or only a dot on the left. One often has to experiment quite a lot to get a good stable trace with certain types of signal. Remember to check that the source of the trigger is the same as the signal you are actually looking at and trying to stabilize. Sometimes the trigger level has to be set carefully at one particular position to give a good stable trace, so perserverance is needed.
- **Trigger Slope:** One can choose whether the sweep will start with a positive or negative going part of the signal. Note that a trigger level alone is not enough to uniquely determine a position on a repetitive signal. The slope at the point must be set as well.

3.2.0 Error Analysis

The accuracy of digital meters is generally better than that of typical analog meters, but there is always some uncertainty in the calibration of any instrument. Good quality multimeters will have a statement about their accuracy on the instrument, typically on the bottom.

3.2.1 Oscilloscope

The voltage scale and time scale of the oscilloscope both have a calibration uncertainty of 3% (for this lab). There is also a reading error of 4% (half of the smallest division on the screen). Combining these two uncertainties in quadrature results in a measurement error for the oscilloscope of 5%.

This error must be applied for both time measurements and voltage measurements. The procedure to be followed to compute the uncertainty in any reading made with the oscilloscope is to add an error of 5% of full scale (not the reading!). This is equivalent to one-half of a major division (or 2.5 times a minor division) for the time scale. Note that this uncertainty is much higher than that of the multimeter and constant for a given scale setting.

The digital frequency readout on the function generator has an associated uncertainty of ± 1 digit.

3.3.0 Experimental Procedure

Step 1: Set up the circuit shown in Figure 3.2. Make sure the power supply is initially set to zero volts (voltage knob fully counter-clockwise) and the current knob is turned fully clockwise.

- Step 2: Turn on both the oscilloscope and the multimeter. Set up the multimeter to measure DC voltages. Select ranges on both the scope and the multimeter for a maximum value you can comfortably observe on the screen. Set the scope to DC coupling so that DC levels can be observed.
- Step 3: For at least four different DC voltage levels, measure the output of the power supply with both the multimeter and the oscilloscope. Record your values in Excel.
- Step 4: Set up the circuit shown in Figure 3.3. Connect the multimeter as a voltmeter. Select the sine function on the generator and a frequency of about 1.0 kHz. To obtain a stationary trace on the scope you will have to adjust both your time base and the frequency. A single period should almost fill the whole screen. Make sure the time base and the vertical gain are both in their calibrated positions.
- Step 5: Take a photo of the oscilloscope screen and record all relevant measurements.
- Step 6: For at least four frequencies between 100 Hz and 1 MHz, measure the frequency of the function generator with both the oscilloscope and the function generator panel meter.
- Step 7: The amplitude of the sinusoidal wave can be measured with both the multimeter and the oscilloscope. Configure the multimeter for measuring AC voltage. The oscilloscope is naturally suited for measuring the peak-to-peak amplitude of a signal. From this the peak, average, and RMS amplitudes can be derived. The point of this step is to deduce which kind of amplitude the multimeter reports. For at least four different amplitudes, measure the amplitude of the sine wave using both the oscilloscope and multimeter. Each of the four sine waves should have a different frequency as well. For best results keep the frequency between 100 Hz and 20 kHz.
- Step 8: For four different amplitudes of a 1.0 kHz square wave, measure the peak and RMS voltage of the square wave, using the oscilloscope and multimeter.

Note that the DC offset control on the function generator adds a DC level to the output waveform. This suggests that readings can be compared between the multimeter and the oscilloscope for the total RMS voltage of a signal with both DC and AC components.

- Step 9: For four different amplitudes and DC offsets of 5.0 kHz sine wave, measure the total RMS voltage using both the multimeter and the oscilloscope. It will be necessary to make two measurements with each instrument, one for the DC component and one for the AC component.
- Step 10: Oscilloscopes can respond to much higher frequencies than multimeters can. Measure the voltage of a sinusoidal wave with a multimeter and then increase the frequency until the multimeter has difficulty reporting a voltage.

Chapter 4

Traveling Waves



Figure 4.1: A photograph of the experimental setup.

4.1.0 Background

Imagine a rope fixed to a wall and under tension at the free end. When this free end is flipped upwards once a wave travels down the rope until it reaches the wall. Here, the wave exerts a force on the wall and the wall exerts an equal and opposite force on the rope. The result is a wave opposite to the original wave traveling back on the rope in the reverse direction. The incident wave has been reflected by the wall. As a second example, suppose now that one end of the rope is free to move up and down because it has been connected to a ring that is free to slide on a frictionless rod perpendicular to the rope. Now, when the free end is flipped up once, a wave travels down the rope as before. When this wave reaches the ring the force of the wave raises the ring which then falls back down to generate a wave travelling back down the rope. Unlike the solid wall, this wave is identical to the original wave, instead of opposite.

These two examples demonstrate that a wave on a rope is reflected by both a fixed rope termination as well as a free rope termination. The fixed and free rope ends are the extreme cases. In an intermediate case, where the rope termination is partly free (or partly fixed), the incident wave will be reflected back with varying intensities and can be opposite or identical, depending on whether the termination is closer to a free end or a fixed end. This logically implies there must be one termination exactly in between which produces no reflection. This proper termination occurs when the rope end has a resistance to the wave that is exactly the same as the

resistance of the rope itself. From the point of view of the rope, the wave continues onward without interruption because the wave only sees more identical rope ahead. From the point of view of an external observer, the proper termination has absorbed the wave and produced no reflection. This single value of resistance that absorbs waves instead of reflecting them is called the **characteristic impedance**, Z_0 . If the rope were thinner or more compliant the characteristic impedance would change.

A second wave property of the rope is the velocity at which the wave propagates along the rope. Giving a harder flip will increase the amplitude of the wave but it will not travel down the rope any faster. This **velocity of propagation**, V_p , is dependent on the thickness and compliance of the rope. One can image reversing the rope so the formerly terminated end is now the end to be flipped. The waves will propagate and reflect on the rope the same as before. For the rope then, the characteristic impedance and velocity of propagation are identical for the forward traveling waves and the reverse traveling reflected waves.

What happens if the rope is flipped twice so that the reflected wave from the first flip must cross the incoming wave of the second flip? The result is determined by a third property of the rope called the **principle of superposition**. This principle states that when two waves overlap, the resulting vertical displacement of the rope (the height or amplitude), is the sum of the individual displacements of each wave. The waves themselves do not interfere with each other. After the waves have crossed, each carries on as before without any change in speed or amplitude.

If the rope were ideal, or **lossless**, then the traveling waves would never lose any energy. In practice, a wave traveling on a rope loses a constant fraction of its amplitude per meter of travel along the rope. This **attenuation**, or **loss factor**, of the rope is a fourth characteristic property of the rope. Due to attenuation, a wave traveling on a very long rope will slowly decay away. In this experiment you will be observing and verifying each of these four properties.

The waves on a rope are an example of a mechanical traveling wave. Traveling wave phenomena occur in any situation where the travel time of a wave along the medium is significantly larger than the size of the wave. Examples of this are everywhere. Bats use acoustic traveling waves in air to locate their prey and avoid obstacles. Dolphins and submarines do the same with acoustic waves in water. The wake from a boat generates surface traveling waves on top of the water that take some time to reach the shore. Striking a hammer on a steel beam generates traveling waves in a solid medium. A sudden application and removal of heat generates a heat wave that travels by conduction. The sudden diffusion that occurs when one material is added to another is also a traveling wave. Lastly, electrical signals in wires and electromagnetic waves in space furnish many examples of traveling waves.

The size and speed of a traveling wave system can vary greatly. The yearly thermal cycle on a concrete dam generates waves with widths in the tens of meters. On the other hand, in a high frequency radar set the waves travel millimeters in a few nanoseconds. This experiment will examine electrical pulses in cables several meters long. The waves will travel the length of the cable in several tens of nanoseconds. This system is used for traveling wave experiments because it closely approximates an ideal traveling wave system. Furthermore, it is easy and fast to make direct measurements on the traveling waves and the cables are small and easily handled. This experiment would not be so easy to perform with yearly thermal waves in a concrete dam.

Let x be the position of the wave and t the time. Then all lossless linear traveling wave systems obey the ***telegrapher's equations***

$$\frac{\partial v}{\partial x} = -\frac{Z_0}{V_p} \cdot \frac{\partial i}{\partial t} \quad (4.1.1)$$

and

$$\frac{\partial i}{\partial x} = -\frac{1}{V_p Z_0} \cdot \frac{\partial v}{\partial t} \quad (4.1.2)$$

Since this experiment deals with electrical pulses on a cable, the variable v represents the voltage and the variable i represents the current. For any other traveling wave system they would represent some other relevant physical quantities. The partial derivatives are required because the wave depends on both the position and the time. Hence the voltage, v , is really a function $v(x, t)$, and the current, i , is really a function $i(x, t)$. The solutions to the telegraphers equations are

$$v(x, t) = f(x - V_p t) + g(x + V_p t) \quad (4.1.3)$$

and

$$i(x, t) = \frac{1}{Z_0} [f(x - V_p t) - g(x + V_p t)] \quad (4.1.4)$$

The solution functions f and g are arbitrary. This means it does not matter what shape the traveling wave may have, the general behaviour remains unchanged. Both the f and g functions each satisfy the telegrapher's equations alone. The reason for this is that f and g are arbitrary, so either f or g could be identically zero. This allows us to examine f and g independently. For f , as long as x is increased with velocity, V_p , the value of $f(x, t)$ remains unchanged. So f is a wave traveling in the positive x direction. Similarly, g is a wave traveling in the negative x direction. Set g to zero so the forward wave can be examined. Dividing Equation 4.1.3 by Equation 4.1.4 gives

$$\frac{v(x, t)}{i(x, t)} = Z_0 \quad (4.1.5)$$

which is a travelling wave version of Ohm's law. For the reverse wave g , the result is

$$\frac{v(x, t)}{i(x, t)} = -Z_0 \quad (4.1.6)$$

It is now possible to describe reflections quantitatively. Imagine a wave traveling on a cable that is terminated in a load Z_T and has a characteristic impedance Z_0 . Let v^+ be the voltage traveling toward the load and v^- be the voltage reflected from the load. The voltage at the load (terminated end) must be

$$v_T = v^+ + v^- \quad (4.1.7)$$

and using Equations 4.1.5 and 4.1.6, the current must be

$$i_T = i^+ + i^- = \frac{1}{Z_0} (v^+ - v^-) \quad (4.1.8)$$

The voltages and currents at the terminated end are determined by the load Z_T , which means

$$\frac{v_T}{i_T} = Z_T = \frac{v^+ + v^-}{v^+/Z_0 - v^-/Z_0} = Z_0 \frac{v^+ + v^-}{v^+ - v^-} \quad (4.1.9)$$

Define the reflection coefficient to be the ratio of the reflected voltage to the incident voltage

$$\Gamma_r = \frac{v^-}{v^+} \quad (4.1.10)$$

Using Equation 4.1.9 this can be rewritten as

$$\Gamma_r = \frac{Z_T - Z_0}{Z_T + Z_0} \quad (4.1.11)$$

Equation 4.1.11 is our desired result. It predicts what the reflection voltage should be for any type of termination. If the load resistance is equal to the characteristic impedance the reflection coefficient is zero. When the termination is $Z_T = 0$ (a shorted end), the reflection coefficient is -1 . When the termination is infinite (an open end), the reflection coefficient is $+1$. This corresponds exactly to what was discussed in the rope example.

4.2.0 Experimental Procedure

Step 1: Connect the output of the pulse generator to the input of the oscilloscope through the feedthrough termination supplied on the oscilloscope. This termination makes the input of the oscilloscope have the same characteristic impedance as the output of the pulse generator so that reflections are eliminated. Set up the pulse generator for maximum amplitude and minimum pulse width. Adjust the pulse period so that the pulses (waves) are spaced 5 microseconds apart. The remaining pushbuttons on the pulse generator must be in the out position, except for POS and INT LOAD which must be pushed in. The pulse generator is now set up to generate waves that are narrow enough and spaced far enough apart that reflections can easily be observed.

Step 2: Determine the maximum and minimum values of the pulse. Also, measure the time it takes for the signal to rise from minimum to maximum (the rise time), and from maximum to minimum (the fall time). Measure the width of the pulse at the half height point. Make two sketches or take two snapshots of this wave. One diagram should illustrate a single pulse in detail. The other diagram should illustrate what a train of pulses looks like. Include relevant measurements/annotations on the diagrams.

Step 3: Connect the circuit shown in Figure 4.2. Observe the reflected pulse and determine the time between the initial pulse and the reflected pulse. Also, measure the amplitude of the initial and reflected pulses. From these you will be able to determine the velocity of propagation and the attenuation of the wave medium (the RG58 cable). Draw a diagram. Connect a short circuit termination to the end of the cable. Draw a diagram with relevant measurements for this case. Repeat this step for the RG59 and twinaxial cables.

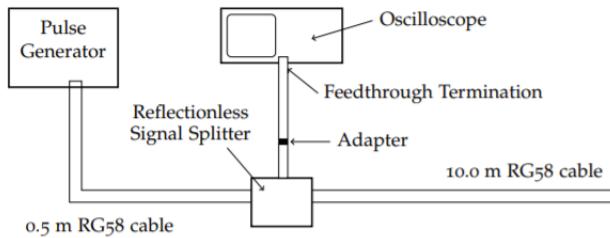


Figure 4.2: A diagram of the experimental set up.

- Step 4: Remove the short circuit termination and connect the variable resistor termination. Adjust the resistor until there is no reflection. Remove the variable termination from the circuit and use the Fluke meter to measure its resistance. This is the value for the characteristic impedance of RG58 cable. Now vary the resistance over its entire range. At each point measure the resistance of the termination and the corresponding amplitude of the initial and reflected waves. Take enough data so that you will be able to verify or refute Equation 4.1.11.
- Step 5: Remove the variable resistance and leave the cable end open. Increase the pulse width until the outgoing pulse overlaps the reflected pulse. Draw a diagram and take amplitude measurements that will allow you to verify or refute the principle of superposition. Decrease the pulse width back to minimum.
- Step 6: Connect the 7.0 m RG59 cable to the end of the RG58 cable. Terminate the open end with a short. Sketch what you observe. Describe what each reflected wave represents. Apply the variable resistive termination to the end of the RG59 cable to measure its characteristic impedance. Repeat with the 7.0 m twinaxial cable.
- Step 7: Set up the circuit shown in Figure 4.3. Connect one end of the 10.0 m RG58 cable to the t-splitter (connected to channel 1) and the other end to channel 2 of the oscilloscope via the feedthrough termination. Set the oscilloscope display to show both channels 1 and 2 and record the voltage differences between the direct output of the pulse generator and the attenuated signal from the cable. Next, measure the time separation between each pulse to determine the phase velocity of the cable. Repeat with the RG59 and twinaxial cables.

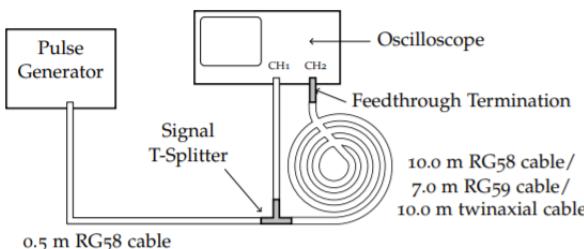


Figure 4.3: A diagram of the experimental set up for the determination of signal attenuation in RG58, RG59, and twinaxial cables.

4.3.0 Error Analysis

The uncertainty in measuring the voltage and the time of the wave is the instrument error of the oscilloscope. Assume that the scale on the oscilloscope screen is a ruler. Then the error is one half of the smallest division. Each division in the horizontal direction will have a size determined by the horizontal oscilloscope setting. Similarly, a vertical division is determined by the vertical oscilloscope setting. Take the length of each cable to be exact. For the multimeter, the error is the instrument error as written on the underside of the meter.

Chapter 5

X-Ray Diffraction

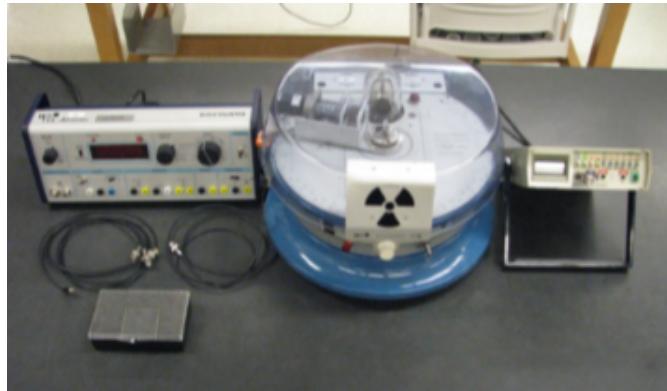


Figure 5.1: A photograph of the experimental setup.

5.1.0 Background

In the late 19th century physicists and inventors began to work with a new device called a cathode ray tube. A cathode ray tube is a glass bulb with the air removed, a hot metal filament at one side, a metallic plate at the other end, and a high voltage between the plate and the filament. This device simultaneously started the eras of modern quantum physics and modern electronics. Among its many other properties, it was noticed that operating such a cathode ray tube at high voltages had a side effect. The cathode ray tube caused various objects outside of the tube to glow. Although other observers had noticed this effect, it was the german physicist Wilhelm Röntgen (1845-1923) who thought it might be important and undertook a systematic study of the phenomenon.

Röntgen noticed that the glow persisted even when the cathode ray tube was covered up. He called the unknown emissions originating from the cathode ray tube x-rays. A name which remains to this day, although they are also known as Röntgen rays in his honor. Moreover, placing a hand between the glow and the tube cast an eerie shadow that illuminated the bones inside his hand in a different light than the surrounding tissue. Röntgen found that this shadow could be photographed. Figure 5.2 shows one of Röntgen's pictures, one of the earliest x-ray photographs ever taken, dating from 1895.



Figure 5.2: An x-ray image of a hand

He also found that the x-rays originated in the cathode ray tube at the location where the internal cathode rays struck the metallic plate at the end of the tube. He then established that their degree of penetration through an element depended upon its atomic number. Elements of high atomic number were able to block the x-rays better. Knowing this, he was then able to use slits cut in lead to show that x-rays travel in straight lines and that they are unaffected by electric and magnetic fields. This suggested that the x-rays were a form of electromagnetic radiation similar to light. For his work on x-rays, Röntgen was given the 1901 Nobel prize in physics, the first ever awarded.

If x-rays were similar to light, it seemed reasonable to check whether they exhibit other optical properties of light such as diffraction. Diffraction is a wave property of light that makes the edges cast by shadows indistinct instead of sharp and can be used to separate light of different wavelengths. Subsequent investigators had difficulty observing any diffraction and began to suspect that if x-rays were electromagnetic waves like light, they must be of a greatly different wavelength. The German physicist Max von Laue (1879-1960) realized that the spacing of atoms in crystals might be suitable for observing the diffraction of x-rays. His insight was successful and he and his students were the first people to observe x-ray diffraction through a crystal. For this work von Laue was awarded the 1914 Nobel prize in physics.

Several more Nobel prizes were earned by physicists studying x-rays. The British son and father team of Sir William L. Bragg (1890-1971) and Sir William H. Bragg (1862-1942) extended von Laue's work. They predicted how x-rays reflect off crystals (a rule now called Bragg's law of x-ray diffraction) and used this to design and construct instruments that can examine x-rays by their wavelength. Such instruments are now called Bragg x-ray spectrometers. For their work the Bragg's received the 1915 Nobel prize in physics. Further work by the British physicist Charles Barkla (1877-1944) established how x-rays are scattered, polarized, and absorbed. He also discovered that when one of its electrons transitions to higher energy level each element emits its own characteristic wavelengths of x-rays which he named K-shell x-rays. For this work Barkla received the 1917 Nobel prize in physics. More results on x-ray atomic spectra and still greater refinement of x-ray spectrometers earned the Swedish physicist Karl Siegbahn

(1886-1978) a Nobel prize in 1924.

It is now accepted that x-rays are a relatively high energy form of electromagnetic radiation with a much shorter wavelength than light. Ultimately, x-rays became a standard diagnostic tool universally used by astronomers, biologists, chemists, dentists, doctors, engineers, geologists, and physicists. The penetrating nature of x-rays makes it possible to image the interior of living organisms, observe cracks hidden in structures and mechanisms, discern the behaviour of energetic astronomical objects, study the arrangement of atoms in crystals, powders, and alloys, determine the composition of surfaces, and probe the internal energy levels of atomic isotopes.

A point most of these x-ray applications have in common is that they require an x-ray generator, a source of x-rays. Aside from refinements, the **x-ray tubes** used in modern x-ray generators (and the x-ray tube used in this experiment) have the same structure as the original x-ray tubes Röntgen used. Figure 5.3 shows the structure of an x-ray tube. A hot **cathode** inside an evacuated glass tube emits electrons. These electrons are accelerated to the other end of the tube by a high voltage and strike the metal target, called the **anode**. This **electron beam** has enough energy to directly interact with the atoms in the target anode. Most of this energy (about 95%) is lost in the target as heat. The remainder of the energy is converted into x-rays. The target is **bevelled** to give some directionality to the emitted x-rays. Further control of the x-ray beam path is achieved by a lead glass envelope that blocks x-rays in all directions, except for the **aperture** through which the x-rays are permitted to pass.

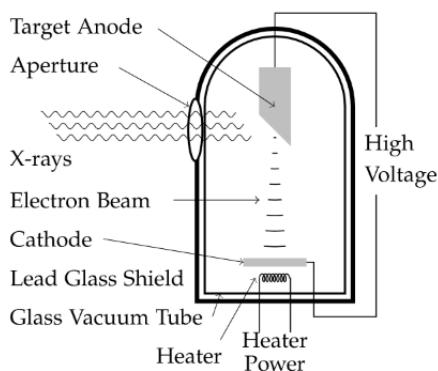


Figure 5.3: Schematic representation of an x-ray tube.

The x-rays generated by a typical x-ray tube cover a broad range of wavelengths. The proportion of x-rays at any given wavelength is governed by the physical interaction processes between the electrons in the electron beam and the atoms inside the target. A typical spectrum of x-rays emitted by an x-ray tube is shown in Figure 5.4. It can be seen that the spectrum exhibits two main features. There is a broad curve at all wavelengths to the right of a certain value. Superimposed on this is a set of sharp peaks. It has been found that the shape of the smooth part is determined by the magnitude of the high voltage and the aperture material. The positions and relative sizes of the sharp peaks are determined by the type of atoms the target anode is composed of.

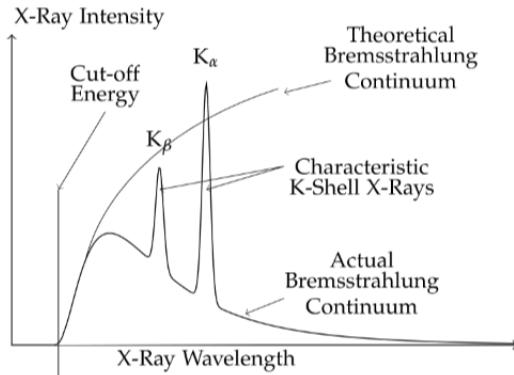


Figure 5.4: An example of an x-ray spectrum.

The smooth broad curve is known as the **bremsstrahlung continuum**. This bremsstrahlung radiation, or **braking radiation**, arises due to the acceleration the electrons undergo when they strike the target anode and interact with the electric fields of the atoms inside. Electromagnetic theory predicts that accelerated charges will radiate electromagnetic waves. The large number of electrons in the electron beam are all accelerating in different ways since they interact with different atoms of the target at a variety of distances. So the radiated electromagnetic energy happens at many wavelengths and this yields the continuum.

According to the quantum picture, this radiation will be in the form of photons with energy $E = hf$ corresponding to the change in the electron kinetic energy ΔK , so that

$$E = hf = \Delta K \quad (5.1.1)$$

where f is the frequency of the photon, and $h = 6.62606876 \times 10^{-34} \text{ Js}$ is **Planck's constant**.

The most energetic possible photon occurs when the electron loses all of its kinetic energy in a single interaction. For an x-ray tube with accelerating voltage, V , the maximum photon energy, E_{max} , and corresponding minimum wavelength λ_{min} is

$$E_{max} = hf_{max} = \frac{hc}{\lambda_{min}} = eV \quad (5.1.2)$$

where $e = 1.60217646 \times 10^{-19} \text{ C}$ is the charge of an electron and $c = 299792458 \text{ m/s}$ is the speed of light. This highest energy is known as the **cut-off energy** of the x-ray tube. An x-ray tube cannot generate x-rays with higher energy than the cut-off energy. At the low end of the spectrum, it seems reasonable to expect that low energy x-rays would be produced in much greater quantities than high energy x-rays near the cut-off. Theoretically, this is indeed correct as can be seen by the theoretical curve shown in Figure 5.4. However, in practice, it turns out that the low energy x-rays are not detected in great numbers. The reason for this is that low energy x-rays are easily absorbed. The lowest energy x-rays do not even make it out of the target material. Still more are blocked by the glass walls of the x-ray tube, the air, and the entrance walls of the x-ray detector. This results in the characteristic shape of the bremsstrahlung spectrum seen in Figure 5.4, where numbers of x-rays are small at low energies, increase to some maximum, and then decrease back down to zero at the cut-off.

The other feature of the x-ray tube spectrum is seen to be the sharp peaks that rise above the bremsstrahlung continuum. These peaks are also due to interactions between the electrons in

the electron beam and the atoms in the target. The energy level model of the atom implies that the electrons surrounding the nucleus of an atom have different energy levels called ***shells***. Each electron shell has a designated ***principal quantum number***, n . As seen in Figure 5.5, the innermost shell has quantum number $n = 1$ and is commonly called the ***K-shell***. The next shell outward has quantum number $n = 2$ and so on. Any transition of an electron from one level to a lower level must be accompanied by the emission of a photon of energy equal to the energy difference between the two levels.

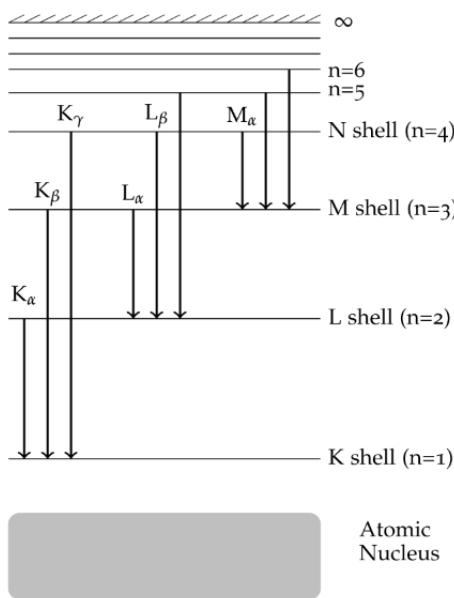


Figure 5.5: Diagram representing electronic energy levels in an atom.

Electrons in the electron beam of the x-ray tube which are sufficiently energetic can penetrate into the atoms and knock an electron out of an inner shell. In turn, an electron from a higher energy level falls down to fill the hole in the inner shell and thereby releases a photon of the corresponding energy difference. So, for example, when an inner electron from the K-shell is removed from the atom, a common transition is for an electron in the next higher shell ($n = 2$) to move into the inner shell and emit a photon. As seen in Figure 5.5, this photon is named K_α . It turns out that for many materials, the K_α photon (and many others) happens to have sufficient energy to be an x-ray photon. Each element has a characteristic set of x-ray peaks corresponding to transitions of electrons between inner energy levels. The height of each peak is related to the probability that the particular electron transition corresponding to that energy difference can happen. These ***characteristic x-rays*** serve as an identifying marker that can be used to identify the target material and explore its inner atomic energy levels.

It is therefore of strong interest to find a method to determine the wavelengths of the features seen in the x-ray spectrum because the actual wavelengths could be used to measure the energy levels of electronic orbitals inside the atom. What is needed is a device that can separate x-rays by their wavelength. For visible light, the device commonly used for this purpose is called a ***diffraction grating***. It consists of a large number of equally spaced parallel lines inscribed onto glass. The spacing of the lines is critical to the operation of the grating. For useable results, the separation between the lines should be roughly the same size or smaller than the wavelength of the light to be diffracted. So for visible light with a wavelength around 500 nm,

the spacing of the lines needs to be at least 200 lines per millimeter. Unfortunately, it turns out that the wavelength of x-rays is about 1000 times smaller than the wavelength of visible light. It is quite difficult to machine lines fine enough to serve as a diffraction grating for x-rays. However, the spacing between atoms in a regular solid such as a crystal serves very well as a natural diffraction grating for x-rays.

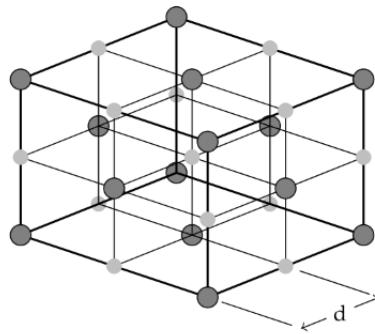


Figure 5.6: Crystalline structure of Lithium Flouride.

This natural correspondence between the x-ray wavelength and the spacing of atoms enables crystals to be used as diffraction gratings for x-rays. The study of crystals and x-rays constitutes the scientific field of ***x-ray crystallography***. There is a large abundance of different types of crystals, each with their own structure and properties. This experiment restricts its attention to the simplest type of crystal, an ionic solid with a cubic structure as seen in Figure 5.6. The most common example of this type of solid is ordinary table salt, NaCl. Unfortunately, salt crystals have a tendency to absorb water and break down. For this reason, the cubic ionic crystal used in this experiment instead is Lithium Fluoride, LiF. As seen in Figure 5.7, each atom of lithium is surrounded by six fluorine atoms and conversely, each fluorine atom is surrounded by six lithium atoms. All of the atoms are the same distance from each other and this gives rise to the cubic crystal structure

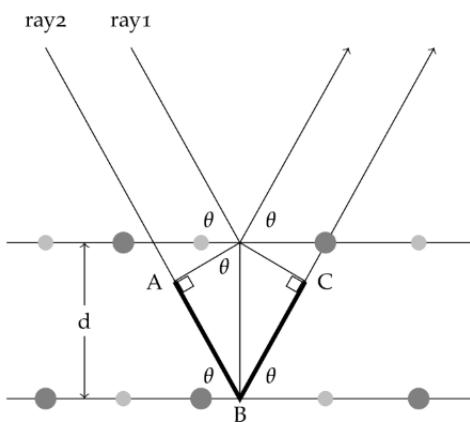


Figure 5.7: A geometrical perspective on x-ray interaction with a LiF substrate.

It seems reasonable to expect that the spacing between the atoms in the LiF crystal influences how the x-rays diffract through the crystal in the same way that the spacing between ruled lines

determines how visible light diffracts through an optical diffraction grating. The cubic structure of LiF makes it possible to calculate the distance between the atoms. Let the atomic mass of lithium be denoted by A_{Li} and the atomic mass of fluorine be denoted by A_F . Then the molar mass, A_{LiF} , of lithium fluoride is given by

$$A_{LiF} = A_{Li} + A_F \quad (5.1.3)$$

The number of molecules in A_{LiF} (one mole) grams of LiF is N_A , where $N_A = 6.02214199 \times 10^{23}$ is Avogadro's number. If the density of LiF is ρ grams per cubic centimeter, then the volume, V , of N_A molecules of LiF must be

$$V = \frac{A_{LiF}}{\rho} \quad (5.1.4)$$

cubic centimeters. Since LiF is diatomic, the atomic spacing, d , in centimeters between the atoms in LiF is

$$d = \sqrt[3]{\frac{A_{LiF}}{2N_A\rho}} \quad (5.1.5)$$

To calculate how the diffraction of x-rays happens inside an LiF crystal, assume that the LiF has a cubic structure and that the x-rays actually are electromagnetic waves. The separation, d , between two neighboring atomic planes is then given by Equation 5.1.5. Imagine two x-rays, ray1 and ray2, with wavelength, λ , reflecting off of two planes of atoms inside the crystal with angle of incidence θ as seen in Figure 5.7. For electromagnetic waves, the angle of reflection equals the angle of incidence, so the reflected ray also leaves the crystal at the same angle θ . If ray1 and ray2 are parallel, then the extra distance travelled by ray2 is the length of the two line segments AB and BC . For constructive interference to occur, the distance $AB + BC$ must be an integer multiple of the wavelength. So when x-rays reflect off of the surface of a LiF crystal, each wavelength reflects at a particular angle and all the other wavelengths cancel out. The length of AB is $d \cdot \sin(\theta)$, so the condition for constructive interference is

$$n\lambda = 2d \sin(\theta) \quad n = 1, 2, 3, \dots \quad (5.1.6)$$

This formula is known as Braggs law and it describes how x-rays diffract when reflected by a cubic crystal. This diffraction interaction is called Bragg x-ray diffraction and it implies that a LiF crystal can be used to measure the wavelength of x-rays simply by tilting it at different angles to the x-ray beam.

The instrument that measures an x-ray spectrum by taking advantage of Bragg's law is called a Bragg x-ray spectrometer. The construction of a typical Bragg x-ray spectrometer is shown in Figure 5.8.

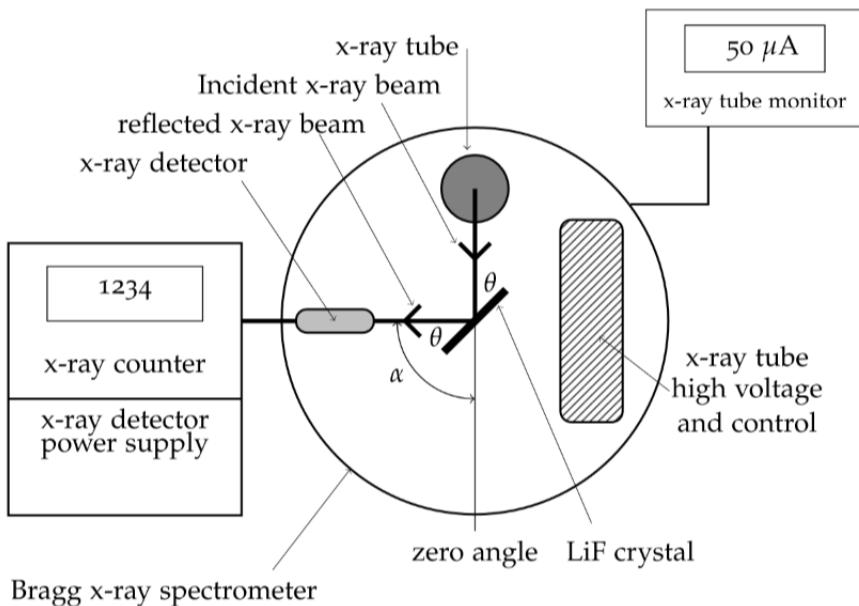


Figure 5.8: An example of a typical Bragg x-ray spectrometer.

An x-ray tube generates a beam of x-rays. The x-ray tube is operated by an associated high voltage power supply together with control and monitoring electronics. The x-ray beam strikes a mounted crystal at some incident angle θ . Bragg diffraction suggests that the reflected x-rays at the same angle θ will be at the single x-ray wavelength given by Bragg's law, Equation 5.1.6. The reflected beam is detected by an **x-ray detector**. An x-ray detector measures the intensity of the reflected beam. Usually this is done by counting the number of arriving x-ray photons. The x-ray detector also requires its own power supply and control system.

The apparatus is arranged so that the detector can be rotated to any angle α . An internal mechanism links the rotation of the detector to the rotation of the crystal. This is done so that for any detector angle, α , the crystal is positioned such that the angle of incidence, θ , equals the angle of reflection, θ , as seen in Figure 5.8. The angle α is always twice the angle θ so that

$$\alpha = 2\theta \quad (5.1.7)$$

In this way the number of x-rays at any wavelength is determined and when the resulting data is plotted, an x-ray spectrum such as the one in Figure 5.4 is obtained. The crystal spacing and Equation 5.1.6 can then be used to relate an x-ray wavelength to each angle. Equation 5.1.6 also suggests that higher order spectral features may be visible when $n = 2, 3, \dots$. In general, higher order spectra are less bright (a smaller number of x-ray photons) than the primary ($n = 1$) spectrum because the extra travel distance through the crystal increases the probability that the x-ray photon will be absorbed inside the crystal. A more detailed view of the Bragg spectrometer actually used in this experiment is shown in Figure 5.9. The most significant addition to the instrument are **safety interlocks and shielding**. Due to their penetrating nature, x-rays are considered to be harmful radiation when received in large doses. For this reason, practical x-ray spectrometers are shielded and are equipped with interlocks so that it is impossible to operate the instrument when the shielding is removed. The first layer of shielding is a **dome shield** completely surrounding the x-ray tube and made of leaded glass. An **aperture slit** is mounted in this shield so that a narrow beam of x-rays can escape and strike the diffraction crystal mounted

at the center of the x-ray spectrometer. The second layer of shielding are the two lead slits (the ***entrance slit*** and ***collimating slit***) that block the diffracted xrays from escaping the instrument, except for a narrow beam that is permitted to reach the detector. The angular resolution of the x-ray spectrometer is determined by the width of these two slits and the distance between them. The third layer of shielding is the ***detector tube*** and ***detector socket*** assembly which absorbs most of the remaining x-rays in the diffracted beam. The fourth layer of shielding is the cover of the instrument. The spectrometer interlocks are designed so that the x-ray tube cannot be energized unless the cover is closed and properly centered. The fifth and final layer of shielding is the ***aluminum and lead backstop*** that blocks the remaining portion of the direct non-diffracted portion of the x-ray beam. Under normal operation with these safeguards in place, the x-ray spectrometer used in this experiment has been certified as radiologically safe for educational use.

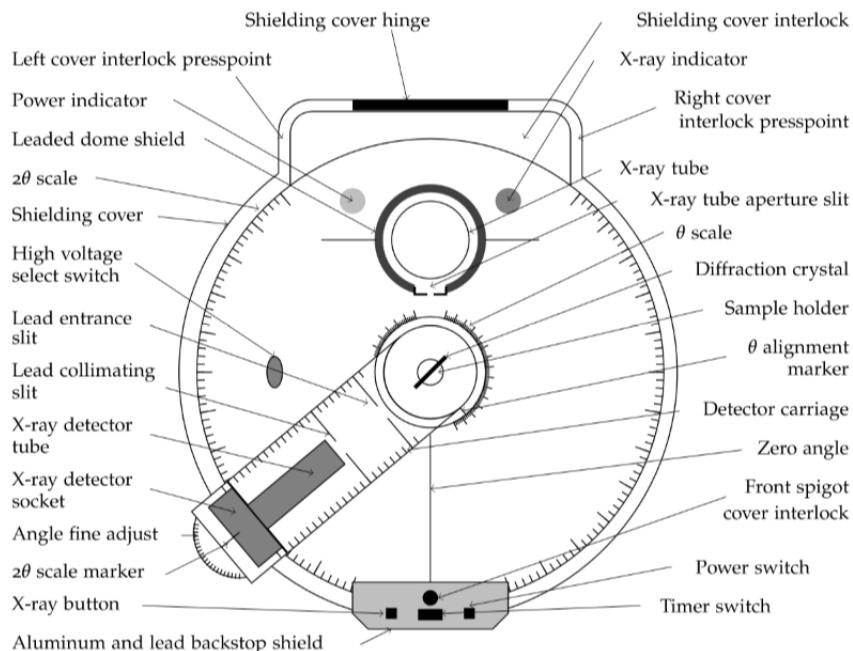


Figure 5.9: A detailed view of the Bragg spectrometer used in this experiment.

The x-ray spectrometer shown in Figures 5.8 and 5.9 can be used to experimentally examine both the production of x-rays by the x-ray tube and the diffraction of x-rays by the sample crystal mounted in the instrument. The x-ray tube used in this experiment has a copper target anode. So it is expected that the K-shell peaks visible in the spectrum should originate from copper atoms and have the characteristic energies of copper. The x-ray tube is operated at a voltage of 30 kV and would be expected to produce a bremsstrahlung spectrum with a cut-off as specified by Equation 5.1.2 if the theory of braking radiation is correct. The crystal used to diffract the x-rays is lithium fluoride. The density of LiF can be used to deduce the atomic spacing of this crystal assuming a cubic structure. The Bragg spectrometer can then be used to measure the spectrum emitted by the x-ray tube by moving the detector to different angles and recording the x-ray photon count at each angle. Bragg's law relates the detector angle to the x-ray wavelength assuming the atomic structure of the crystal is understood. Analysis of the x-ray spectrum allows to test Bragg's law and the theory of x-ray production and measurement. Lastly, bremsstrahlung can be further examined by placing a thin piece of plastic in the x-ray

beam. If the reason for the long low energy tail is due to x-ray absorption, then a piece of plastic should attenuate the low energy part of the spectrum to a much greater extent than the high energy portion. This can be checked by comparing x-ray spectra with and without the plastic against each other.

As is common in spectroscopy, there is a dilemma between accepting the properties of the diffraction grating (in this case a crystal) or accepting the properties of the light source (in this case an x-ray tube). If the grating is trusted, then the grating can be used to measure the spectrum of the incoming x-rays. If the x-ray tube is trusted, then the x-ray beam can be used to examine the atomic structure of the sample being irradiated. In this experiment, neither device is treated as the accepted standard. Instead, the accepted values for the atomic spacing of LiF, the K_α and K_β wavelengths of copper, and the bremsstrahlung cutoff are compared with the measured values to establish the consistency of the x-ray spectrometer system with x-ray theory. In practice independent evidence is gathered to establish the behaviour of x-ray spectrometers. The structure of LiF is independently deduced by the known chemical properties of ionic solids and the fracture properties of LiF crystals. The wavelengths of x-rays are independently verified without crystals by using ruled gratings that are tilted nearly horizontally to effectively decrease the line spacing of the grating to the point where x-ray diffraction can be obtained.

5.2.0 Experimental Procedure

Step 1: The apparatus is connected together as shown in Figure 5.8. Make sure all instruments are turned off before proceeding. The x-ray spectrometer is a nearly self-contained unit except for the x-ray counter and the tube current monitor. The x-ray detector cable should be connected to the GM tube input connector on the right on the Tel-Atomic Digicounter. The monitor tube current output should be connected to the current jacks on the digital multimeter with the supplied cable. Set the multimeter to a range compatible with a current on the order of $50 \mu\text{A}$ DC.

Step 2: The Tel-Atomic Digicounter powers the x-ray detector tube and accumulates and displays the detected x-ray photon counts. For consistent repeatable operation, that enables comparisons between different instruments, the following settings are recommended. The GM Tube Supply voltage should be 420 V. The counting switch should be set to continuous and the function switch turned to radioactivity. Adjust the range switch to 10 s. Set the timing switch to stop and the Triggered setting to off. The remaining switches and connections can be neglected.

Step 3: The following checks and configurations are recommended for proper operation of the x-ray spectrometer. The names of the various spectrometer parts as used here are shown in Figure 5.9. First lift the spectrometer cover (shielding cover). This cover has an interlock, so there is a special procedure for opening it. As seen in Figure 5.9, there are cover interlock presspoints on both sides of the spectrometer at the back. To open the cover, press the right interlock presspoint to shift the entire cover left. The cover can be opened once the cover is shifted fully leftwards. The cover hinge is at the rear, so it is opened by reaching under the backstop shield and raising the opening from the front. Once the cover is lifted, examine the instrument. Do ***NOT*** touch the crystal, it is brittle and sensitive to moisture, but check that it is present,

centered in the sample holder, and the top of it has a dab of blue paint. The blue paint identifies the crystal as LiF. Make sure the x-ray tube is installed and fully covered by the leadglass dome shield. Inspect the backstop shield to make sure there is a thin plate of lead behind the plate of aluminum on the outside of the cover. Set the high voltage select switch to 30 kV. Check that a 1 mm lead aperture slit is installed at the exit port of the x-ray tube. Note that the carriage holding the detector has numbered slots. A 1 mm lead collimator should be present in slot 13 of the detector carriage and a 3 mm lead collimator should be installed in slot 18. Lastly, the x-ray detector itself should be installed in slot 26 of the detector carriage.

- Step 4: Check that the detector carriage moves freely between 20° and 120° on the 2θ scale. For proper operation, the detector carriage must always be on the left side of the instrument as seen in Figures 5.8 and 5.9. The reason for this is that the crystal has a preferred face where it has been carefully cleaved, and this is the side that should be exposed to the x-ray beam. Verify that the mechanism which rotates the crystal is working correctly. To do this, rotate the detector and, if possible, compare the reading on the 2θ scale with the reading on the θ scale. One should be double the other. The 2θ scale is read by the scale marker just behind and underneath the detector socket. The θ scale is read by the angle alignment marker line that is visible on the ring surrounding the sample holder. Next, check the zero alignment of the spectrometer. This is done by rotating the detector to 0° on the 2θ scale. When the zero alignment is correct, the θ scale should also be registering 0° at this point. If there is a misalignment, contact the laboratory staff
- Step 5: At this point the instruments are ready to be turned on and the interlocks on the spectrometer checked. Lower the lid and center it using the left and right interlock presspoints. The cover is correctly centered when the front spigot interlock is centered on the the zero angle line. Check (gently) that the cover cannot be opened when it is centered at this position. Turn on the power to the Digicounter and the multimeter. To power up the spectrometer, rotate the timer switch clockwise to the 50 minute mark and then turn on the power switch. The power indicator should light up and the filament inside the x-ray tube should begin to glow. Then press the x-ray button. If the cover is correctly centered, the x-ray indicator should light up and the spectrometer will begin generating x-rays. If the x-ray indicator does not light, it usually implies that the cover needs a minor position readjustment. Once turned on, check the interlock by pressing the right interlock presspoint and shift the lid left. The x-ray indicator should immediately turn off if the interlocks are operating properly. Re-center the cover and turn the x-rays back on. Report to your laboratory instructor and the laboratory staff if the interlock system is not working correctly.
- Step 6: Let the x-ray system warmup and stabilize for about fifteen minutes. Monitor the beam current on the multimeter during this time. The accepted range of operation for the x-ray tube is a beam current between 30 and 70 μA .
- Step 7: The spectrometer is now operational. Obtain an x-ray spectrum. At each degree for 2θ between 20° and 120° record the ten second x-ray photon counts. The Digicounter automatically repeats ten second measurements so the reset button on the Digicounter is not required. It is sufficient to make sure that a full ten second measurement is made at each angle before moving the detector onto the next angle. Occasionally check the timer switch, and when it nears zero, again turn the time clockwise

back up to 50 minutes. For best results keep the spectrometer running continuously for the entire acquisition of the spectrum. The best practice is to adjust the dial to maximum of 50 minutes every 30 minutes, for “fool proof” operation.

Step 8: Insert a plastic plate in a slot between the two collimators on the detector carriage and obtain a second spectrum in the same manner as the first spectrum.

Step 9: Once the second spectrum is completed, turn off the spectrometer and the other instruments.

5.3.0 Error Analysis

It might be expected that an x-ray spectrum obtained with the Bragg spectrometer used in this experiment would have an appearance something like that shown in Figure 5.4. This suggests that the x-ray spectra obtained in this experiment are actually graphs of x-ray photon counts versus angle, where the 2θ detector angle is the independent variable and, N , the number of x-rays, is the dependent variable. As with any graph made from measurements, there will be error bars on each data point reflecting the uncertainty with which it was obtained.

In this experiment, the uncertainty in the angle is dominated by the width of the x-ray beam passing through the instrument. The width of the x-ray beam is determined by the widths and separations of the lead slits in the apparatus. An estimate of the beamwidth can be found by sighting through all of the slits while moving the detector carriage and observing how much the detector must move to each side before the anode of the x-ray tube is blocked. For the instrument used here, it is found that the beamwidth is 3° on the 2θ scale. So the error in 2θ is 1.5° and the error in θ would be at least half of this. Since there are also alignment errors between rotations of the crystal and rotations of the detector, rounding up the error to the nearest degree seems to be prudent. So a reasonable value to use for $u(\theta)$ would be 1° .

Additionally, due to the difficulty in observing angles between the physical lower limit of the spectrometer, $2\theta = 18^\circ$, and the regime of easily measureable angles, $2\theta \geq 20^\circ$, uncertainties in these small-angle measurements should be chosen appropriately.

The detection of x-rays is a statistical process. Each x-ray photon has a particular probability of being counted by the detector. For this reason, the number of x-rays counts varies, even when the angle remains unchanged. If the x-ray count measurement were repeated a large number of times at each angle, a distribution would be obtained, and the width of the distribution would be a good estimate of the error in the x-ray photon count. However, it is known that for photon counters, the distribution typically is gaussian at large count rates and poissonian at small count rates. In either case, a good estimate of the photon count error, $u(N)$, for both of these statistical distributions is

$$u(N) = \sqrt{N} \tag{5.3.1}$$

Other sources of systematic error are also present. The precision with which the atomic spacing, d , is known affects the precision with which the x-ray wavelengths can be calculated from Bragg’s law. Moreover, the purity, quality, and uniformity of the crystal are large factors in the sharpness of the observed features in the x-ray spectrum. If the spacing of the LiF atoms

is not identically cubic due to impurities or imperfections, then the analysis of Bragg x-ray diffraction as presented here does not fully apply. Lastly, the photons counted by the x-ray detector do not all originate from the x-ray tube. Other radioactive sources such as cosmic rays and trace amounts of isotopes in the spectrometer and surroundings contribute as well. These extra sources of photons are called background radiation. For this experiment, errors in the crystal lattice spacing and background radiation can be neglected.

Chapter 6

Spectroscopy



Figure 6.1: Equipment used in the spectroscopy experiment.

6.1.0 Background

6.1.1 The Spectroscope

Spectroscopy is the study of the electromagnetic spectrum and the emission and absorption of electromagnetic radiation. The science of spectroscopy is useful in many areas that require the analysis of light. Astronomers use spectroscopy to determine the composition of stars and galaxies. Much of what is known about the Sun comes from the analysis of its spectrum. One of the first to study the Solar spectrum was Joseph von Fraunhofer. In 1814 he found that the spectrum of the Sun contained fine, dark lines which, are now known to be caused by the absorption of light by gases in the Sun. A pattern of dark lines in a continuous spectrum is called an **absorption** spectrum. In contrast, discrete coloured lines on a dark background is called an **emission** spectrum and is caused by the emission of light by excited atoms.

The spectral lines visible to the human eye vary from violet to red in colour. The most common method of studying spectra is through the use of a spectroscope. This type of instrument enables the user to observe the spectrum of various elements and facilitates the measurement of the wavelengths of each spectral line. Figure 6.2 shows a schematic diagram of the Introductory Student Spectroscope used in this experiment. Light passes through the slit and the collimator, a device producing a parallel beam of rays. It is then bent by the diffraction grating. The telescope is connected to a rotatable arm so that it can be moved to allow the diffracted light to pass through it. A vernier on the telescope is then used to measure the angle at which the incident light is diffracted. Then this angle can be used to find the wavelength of the selected emission line. Encircling the base of the grating table is a 360 degree scale used to make measurements in conjunction with the vernier scale, which is fixed to the telescope arm.

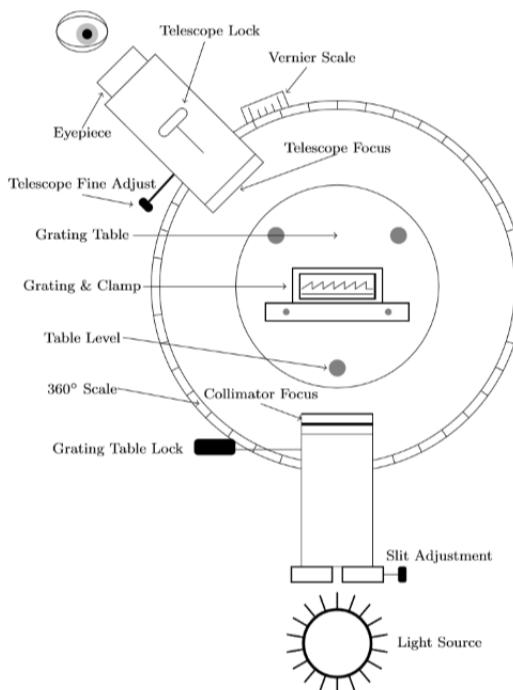


Figure 6.2: A diagram of the spectrometer used in this experiment.

Before the spectroscope can be used to measure the angles of spectral lines it must be adjusted to ensure that accurate results are obtained. To begin, the telescope is focused on a distant object, using the telescope focus. The telescope is then lined up with the slit in the end of the collimator, and the collimator focus is adjusted until the slit comes into sharp focus. In this way only approximately parallel rays of light pass through the diffraction grating.

6.1.2 Diffraction Gratings

A diffraction grating can be imagined as a flat surface with thousands of grooves. As seen in Figure 6.3, the diffraction grating used in this experiment has a sawtooth pattern, called a blazed grating, instead of simple grooves.

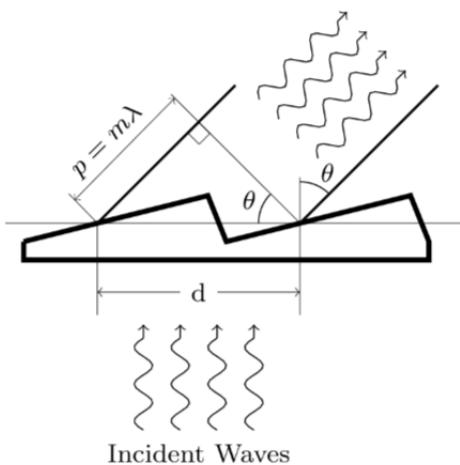


Figure 6.3: Diffraction of incident light by grating.

When light from an emission spectrum passes through a blazed diffraction grating the light is split into its constituent colours. Atomic spectra are observed as a series of lines because of the constructive interference of the light waves once they have passed through the diffraction grating. Destructive interference produces what is observed between spectral lines. Constructive interference occurs when two or more waves of light, of equal wavelength, are superimposed in phase, resulting in a maxima, or a spectral line. In order for two identical light waves, diffracted by side by side grooves, to be in phase they must travel a distance that differs by one wavelength after they have been diffracted. As seen in Figure 6.3, this implies that the distance, p , must be an integer multiple of the wavelength. Therefore, for constructive interference the requirement is that:

$$d \sin \theta = m\lambda, \quad m = 0, \pm 1, \pm 2, \dots \quad (6.1.1)$$

where d is the **line spacing**, that is, the distance from one point on a groove to the same point on the neighboring groove. Note that the number given on the top of the diffraction grating is the **line density**, not the line spacing. Line density is the inverse of line spacing. The **angle of deviation**, θ , is the angle between the central maximum and the spectral line. The variable m indicates the **order** of the spectral lines. First order spectral lines ($m = \pm 1$) are those found closest to the central maximum. As the telescope arm moves away from the central maximum, a point is reached where the spectral lines begin to repeat. These are known as the second order spectral lines ($m = \pm 2$). Although second order maxima are visible, only first order maxima are used in this experiment. The sign of the order is positive or negative depending whether the spectrum is to the right or left of the central maxima. Equation 6.1.1 can be used to calculate the wavelength of a particular spectral line from the angle of deviation. For example, assuming a line spacing of 1000 *lines/mm* and a first order angle of deviation of 30° a wavelength of 500 nm is calculated.

6.1.3 Spectral Sources

The light from the emission spectrum that passes through the diffraction grating is generated by lamps called **spectral sources**. Three different types of spectral sources are used in this experiment: the Geissler tubes, the low pressure lamps, and the gas discharge tubes. In general,

the glass parts of spectral lamps should not be touched. Touch only the metal ends of the tube. After a lamp is turned on move it as little as possible so that its lifetime is not shortened. When finished with a light source allow the lamp to cool for at least five minutes before moving it. Once a tube is removed from the set up, store it away immediately before starting to use a different tube.

The Geissler tubes (like the Hydrogen one) are constructed by filling a glass tube with a pure sample of the desired gas and then adjusting the pressure of the gas to yield maximum brightness. The glass tube is fitted with two electrodes, one at each end. The completed tube is clamped into a high voltage power supply. The high voltage creates an electric field between the electrodes which causes the valence electrons in the atoms to become excited. When the excited electrons lose energy they give off a mixed light containing all wavelengths that are characteristic of the element contained in the tube. Geissler spectral lamps should be turned off when not being used to lengthen their life. Be careful when handling a high voltage supply, do not touch any exposed metal parts.

The low pressure tubes contain a source that is solid at room temperature so there is insufficient vapour present for the lamp to be started at any reasonable voltage (less than several thousand volts). It was found that by adding a noble gas mixture to the lamp (usually Argon and Neon, in various proportions), the starting voltage could be lowered to about 600 Volts. When a cold lamp is turned on a voltage “kick” is supplied to the discharge tube which generates enough charge carriers that the Argon and Neon become conductive. Once conduction begins, the voltage decreases and the discharges remain confined to the noble gases. As the noble gases discharge, heat is generated which eventually causes the source to vaporize. The vaporized atoms are excited by the electric field present in the lamp. The noble gases are not excited any more because the valence electron of the source is much easier to excite than any of the electrons in the noble gases. When using the low pressure lamp it is important to allow the lamp to warm up for at least ten minutes before making any measurements in order to ensure that the spectrum being measured is actually that of Sodium and not of Neon or Argon. The low pressure lamp should be left until all measurements are completed. Turning the lamp on and off will reduce its lifetime, and it is also time consuming as the lamp must be warmed up each time it is turned on.

6.1.4 The Hydrogen Atom

One of the most commonly studied emission spectra is that of Hydrogen. With the simplest spectrum, Hydrogen is the most basic case of emission and absorption to study. Excited atoms can emit electromagnetic radiation at many wavelengths but for the purposes of this experiment, only the visible region of the electromagnetic spectrum is examined. The visible emission spectrum of Hydrogen is known as the Balmer Series, named after Swiss mathematician who first discovered their pattern.

Johann Balmer (1825-1898) was a secondary school mathematics teacher. As an adept mathematician, Balmer was able to produce a formula that predicts the wavelength of Hydrogen spectral lines with great accuracy. His empirical Balmer formula takes the form:

$$\frac{1}{\lambda} = R \left[\frac{1}{2^2} - \frac{1}{n^2} \right], \quad n = 3, 4, 5, \dots \quad (6.1.2)$$

where R is now called **Rydberg's constant** and each integer n corresponds to an observed Hydrogen spectrum line. A proper explanation was not available until the Danish physicist, Niels Bohr (1885-1962) provided his theory of the atom which demonstrated that

$$R = \frac{m_e e^4 Z^2}{8\epsilon_0^2 h^3 c} \quad (6.1.3)$$

where Z is the atomic number of the atom.

6.2.0 Experimental Procedure

- Step 1: Remove the diffraction grating and its holder from the spectroscope. It is extremely important not to touch any of the lenses or the diffraction grating. Do not use fingers or articles of clothing to clean the optical equipment as this will only result in scratching the optics. Hold the diffraction grating by its frame.
- Step 2: The telescope must be focused for parallel light rays and the crosshairs must be put into focus. First slide the eyepiece on the telescope in or out and bring the crosshairs into sharp focus. Next, focus the telescope on a distant object across the room or out a window. (Remember to rotate the telescope to avoid trying to focus through the collimator)
- Step 3: Use the Hydrogen lamp as the calibration light source and place it so that the light from it passes through the slit in the collimator. Aim the telescope at the slit and gently move the lamp until the slit is at its brightest. Adjust the collimator focus until the slit is in sharp focus. In this way the spectrum being measured uses only light parallel to the slit. Align the cross-hairs in the eyepiece with the slit. Do not change the focus of the telescope or the collimator, or rotate the eyepiece for the remainder of the experiment, as this affects the apparent position of the centre slit. Do not move the telescope by touching the eyepiece or the chrome part of the telescope, as this could change the focus of the telescope.
- Step 4: Replace the diffraction grating into its holder. The bolts for the diffraction grating holder should be facing the collimator, and the diffraction grating frame should be placed in the holder so that the diffraction grating is as close to the holder as possible. The idea here is that the diffraction grating needs to pass through the centre of the grating table as seen in Figure 6.2
- Step 5: Turn the telescope and look at spectral lines that are far away from the center angle (choose second order lines if possible). Using the three grating table leveling screws on the bottom of the grating table, adjust the level of the table until the lines are vertically in the middle of the telescope as shown in Figure 6.4. Move the telescope to the other side of centre and adjust the lines until they are also centered. This adjustment may have to be repeated several times on each side.

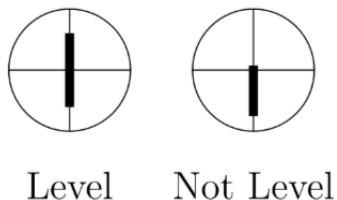


Figure 6.4: Levelling the grating table.

Step 6: Look through the telescope and adjust the width of the slit. Take note of which side of the slit moves. All measurements should be taken from the side of the light line that does not move. By doing this the width of the slit can be changed during the experiment. The slit width adjustment is used to adjust the amount of light that is allowed to pass into the collimator, as different amounts of light are useful in different situations. Some lines are close together and will overlap if the slit is too wide, in which case a narrow slit is required in order to resolve all of the lines that are present. On the other hand some lines are very dim. In such a case a wide slit is better in order for the spectral line to be bright enough that it can be seen.

Step 7: Align the diffraction grating so it is perpendicular to the collimator as seen in Figure 6.3. This is done by rotating the grating table. If the diffraction grating is not perpendicular, the angle of deviation of the spectral lines will differ between the two sides of the centre maximum. Choose any line in the spectrum and rotate the grating until the angle of deviation is the same on the left spectrum as the right spectrum. This step may need to be repeated several times.

Step 8: Measure the angle of the centre maximum. Align the cross-hairs with the stationary edge of the slit. Use the vernier to measure the angle of the maximum. All angles of diffraction are measured with respect to this centre angle. Each angle of diffraction is subtracted from the centre angle and the absolute value of that angle is the angle of deviation. The angle of deviation is then used to determine wavelengths of spectral lines using Equation 6.1.1.

Step 9: Use the Hydrogen spectral lamp to check the performance of the diffraction grating. Measure the angle of deviation on both sides of the maximum for the four lines calculated according to the Equation 6.1.2. The measured angles of deviation, along with Equation 6.1.1, can be used to calibrate the spectroscope. If Equation 6.1.1 holds, a plot of λ versus $\sin \theta$ should yield a straight line with slope d and an intercept of zero.

Step 10: Using wavelengths calculated using Equation 6.1.2 and measured angles of deviation, make a quick plot of λ versus $\sin \theta$ (errors need not be included here). Check that the slope of this line corresponds to approximately 600 *lines/mm*.

Step 11: The spectroscope is now adjusted, checked, and ready for analyzing other spectra. Turn on one of the unknown sources and shine the light through the slit in the collimator. Measure the angles of deviation, on both side of the center maximum, for the visible lines.

Step 12: Repeat these measurements for the other two spectral lamps. Remember, that these lamps have to be warmed up before you can start the measurements.

6.3.0 Error Analysis

Usually $u(\theta)$ would be half the smallest division of the vernier on the spectroscope, however there are other factors that add to the error. First, it is hard to take measurements from the same part of the spectral line each time a measurement is made. Second, as the eye moves back and forth across the eyepiece the cross-hairs move, therefore in order to get precise measurements the eye would have to be in exactly the same place for each measurement. Third, some lines are not very intense making them hard to see, and lining the cross-hairs up becomes difficult. Fourth, the cross-hairs have a thickness which introduces an uncertainty in any visual measurement. Fifth, the condition of the grating, i.e., smudges, scratches, etc., can introduce visual aberrations like smearing. Finally, the spectroscope and diffraction grating used cannot distinguish between spectral lines that are too close together. This is a limitation of the diffraction grating construction as well as the magnification of the telescope. For example, the lines of the Sodium doublet differ by about half a nanometer and they are very close to appearing as one line. For these reasons, take the error in measurements of angles for this spectroscope to be 0.2° . For the uncertainty in the line density of 600 *lines/mm* for the diffraction grating a value of 5 *lines/mm* has been found to be reasonable to use for the diffraction grating check. For other the line density and error determined by the grating performance check can be used. Note that $u(\theta)$ must be in radians when calculations are being made. All other errors should be propagated using an appropriate method.

Chapter 7

Voltage Dividers and Sources

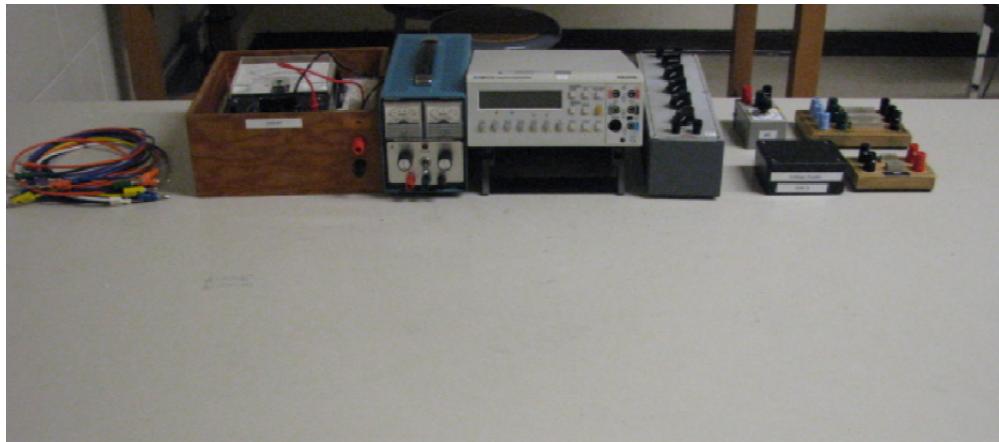


Figure 7.1: Equipment used in the DC Circuit Experiment

7.1.0 Background

In almost all circuits **resistors** are among the most common components. Figure 7.2 shows some of the many different kinds of resistors. The different shapes and sizes play a role in their behaviour and application. Resistors are used to control the current and voltage in a circuit.



Figure 7.2: Examples of resistors of different shapes and sizes.

Georg Ohm (1789-1854) contributed largely to what is now known about resistors. He speculated how current might work and formulated the law governing resistors that now bears his name. Later, Gustav Kirchhoff (1824-1887) made further contributions to understanding of electric circuits. He extended Ohm's work describing what are now called **Kirchhoff's Laws**, which explained how current and voltage in electric circuits are related. In 1883 Léon Thévenin (1857-1926), a french telegraph engineer, described what he thought to be a new theory of equivalent circuits. He showed that any resistor circuit could be simplified to make analysis easier. Coincidentally, the concept of equivalency in circuits had already been proposed almost 30 years earlier by the physicist Hermann von Helmholtz (1821-1894). Thévenin was unaware of Helmholtz's work and both theories met with resistance during their time. It was due to Thévenin's engineering approach, and the growth of electrical engineering in the coming years, that his result is now called Thévenin's Theorem.

Many combinations of resistors exist in circuitry, but the **voltage divider** is one combination that is seen everywhere. The circuit in Figure 7.3 contains a voltage source of some kind, connected in series with two resistors. A voltmeter is in parallel with one of the resistors to measure the voltage across it. Each resistor in the loop drops a portion of the voltage. This resistor combination, which accomplishes the splitting of voltages, is called a voltage divider. It can be used to control the voltages coming out of, or going into, a particular system, or even be used as a tool for analysis of circuits.

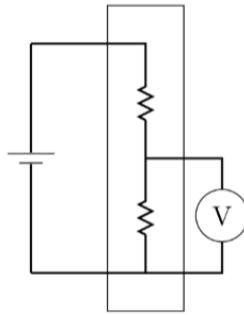


Figure 7.3: Basic voltage divider with power supply.

In Figure 7.4 the connections of the voltage divider to the voltage source and voltmeter are explicitly shown. The voltage source is supplying a voltage with magnitude V_{in} into the voltage divider, and the voltmeter reads V_{out} across R_2 . This suggests the voltage divider can be generalized further by extracting it from circuit

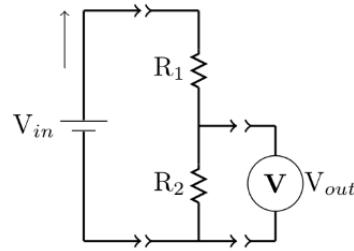


Figure 7.4: Connections for the voltage divider to voltage source and voltmeter.

Figure 7.5 shows only the voltage divider. The input voltage, V_{in} , can be generalized to be any source of voltage such as a battery, power supply or another device. The output voltage, V_{out} , is what would be applied to any device or system connected to these terminals.

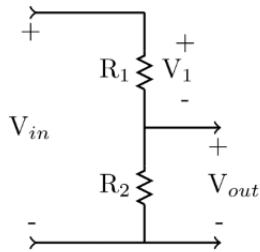


Figure 7.5: Isolated voltage divider.

For the voltage divider in Figure 7.5, V_{in} is split and V_{out} is some portion of V_{in} . This can be shown using Ohm's and Kirchhoff's laws. From Ohm's law, the voltage drop across an ohmic resistor is the product of its resistance and the current through it. So the voltages across R_1 and R_2 are V_1 and V_{out} , which are given by

$$V_1 = IR_1 \quad (7.1.1)$$

and

$$V_{out} = IR_2 \quad (7.1.2)$$

where I is the current flowing in the loop. From Kirchhoff's Voltage Law, it is expected that the applied voltage equals the voltage drops across the resistors in series, so that

$$V_{in} = V_1 + V_{out} \quad (7.1.3)$$

Using Equations 7.1.1, 7.1.2, and 7.1.3, it is seen that

$$V_{in} = I(R_1 + R_2) \quad (7.1.4)$$

Substituting for the current using Equation 7.1.2 gives the voltage divider formula:

$$V_{out} = V_{in} \frac{R_2}{R_1 + R_2} \quad (7.1.5)$$

Equation 7.1.5 implies that the ratio $\frac{V_{out}}{V_{in}}$ is equal to the ratio of $\frac{R_2}{R_1 + R_2}$. It also implies that V_{out} cannot be larger than V_{in} since the ratio of one resistor to two resistors can only be equal to or less than one. A special case is when R_1 is zero. The ratio becomes equal to one and V_{out} is equal to V_{in} ,

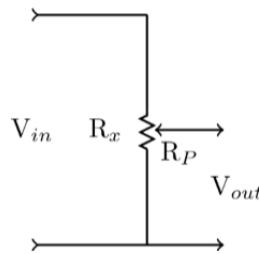


Figure 7.6: Variable voltage divider.

The voltage divider in Figure 7.5 supplies a fixed voltage ratio. It is also possible for voltage dividers to supply a variable ratio. This is done with a device called a potentiometer. A potentiometer is a variable resistor whose resistance can be changed by turning a dial. Figure 7.6 shows a schematic of a potentiometer used as a variable voltage divider. The resistance of the entire potentiometer is R_x and R_p is the portion of this controlled by the dial. If the dial is adjusted so that R_p is equal to R_x , then V_{out} is equal to V_{in} . Likewise V_{out} becomes zero if the dial is adjusted so that R_p is zero.

Suppose two more resistors are added to the potentiometer to get the circuit shown in Figure 7.7. The derivation for Equation 7.1.5 involves a loop with two resistors, but in fact can work for any number of resistors. The total series resistance would be on the bottom of the ratio. The output voltage V_{out} could be taken across any or all of the resistors in the circuit and the value of these resistors would go on the top of the ratio. So, the output voltage will be

$$V_{out} = V_{in} \frac{R_p + R_2}{R_1 + R_2 + R_x} \quad (7.1.6)$$

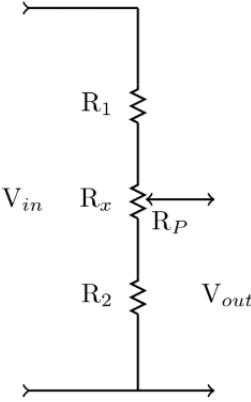


Figure 7.7: Voltage divider with three resistors

This setup produces a range of voltages for V_{out} , which can be controlled by the dial on the potentiometer. The output voltage has a minimum greater than zero and a maximum less than V_{in} , as determined by R_1 and R_2 . The minimum output voltage, V_{min} , occurs when R_P is zero so

$$V_{min} = V_{in} \frac{R_2}{R_1 + R_2 + R_x} \quad (7.1.7)$$

and the maximum output voltage, V_{max} , arises when R_P equals R_x which gives

$$V_{max} = V_{in} \frac{R_x + R_2}{R_1 + R_2 + R_x} \quad (7.1.8)$$

Not only do voltage dividers exist explicitly as the circuits shown in Figures 7.3-7.7, they also exist implicitly whenever any two circuits are connected together. There is a division of voltages between the output of any circuit and the input of a second circuit. For example, imagine that a single resistor is connected to a power supply as shown in Figure 7.8. Here, the second circuit is composed of a single resistor R_L . Typically, R_L is called the load, or in this case the load resistor.

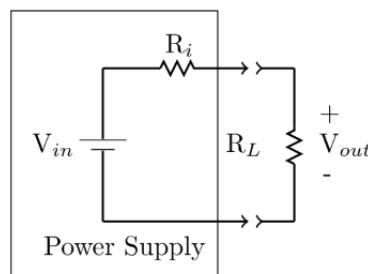


Figure 7.8: Power supply with internal resistor.

Ideally, it would be expected that the entire voltage, V_{in} , would be developed across the load resistor. However, what happens in practice is that the voltage across the load, V_{out} , is always somewhat smaller than V_{in} . Moreover, the smaller the value of R_L , the larger the discrepancy between V_{out} and V_{in} . This situation can be neatly explained by the addition of an internal

resistor, R_i . The voltage is now seen as being split between the load resistor R_L , and the internal resistor, R_i . The basic voltage divider formula, Equation 7.1.5, can be used to find V_{out} as before to get

$$V_{out} = V_{in} \frac{R_L}{R_L + R_i} \quad (7.1.9)$$

What is now observable is that the voltage across R_L is less than the voltage being supplied by the power supply. If R_i is much smaller than R_L , V_{out} is, to a reasonable approximation, equal to V_{in} . At the same time, if the value of R_i is close to R_L , then V_{out} will only be a portion of V_{in} . If R_i were much larger than R_L this effect would be dramatically increased and almost no voltage would be across R_L . For perfect power supplies R_i is zero which means that $V_{out} = V_{in}$ for any load resistor.

$$\frac{1}{V_{out}} = \frac{R_i}{V_{in} R_L} + \frac{1}{V_{in}} \quad (7.1.10)$$

With Equation 7.1.10, R_i can be found by connecting a variable load and measuring V_{out} across different load resistances. It can be seen that plotting $\frac{1}{V_{out}}$ versus $\frac{1}{R_L}$ yields a straight line with slope $\frac{R_i}{V_{in}}$ from which R_i can be found.

Question 7.1.1. What is internal resistance exactly?

Thévenin answered this by showing that from the point of view of any resistor in a circuit, the rest of the circuit is equivalent to a voltage source with a series of internal resistances, just as in Figure 7.7. For example, the left side of Figure 7.9 shows a circuit containing a number of voltage sources and resistors. From the point of view of a single resistor such as R_5 , the entire remaining part of the original circuit is equivalent to a single voltage source in series with a single resistor, as shown on the right side of Figure 7.9. This is called the Thévenin equivalent circuit. The voltage source, V_{th} , is called the Thévenin equivalent voltage and the resistor, R_{th} , is called the Thévenin equivalent resistance. Comparing this with Figure 7.8, it is seen that the Thévenin equivalent circuit is a power supply with an internal resistance driving a load consisting of a chosen component, in this case R_5 .

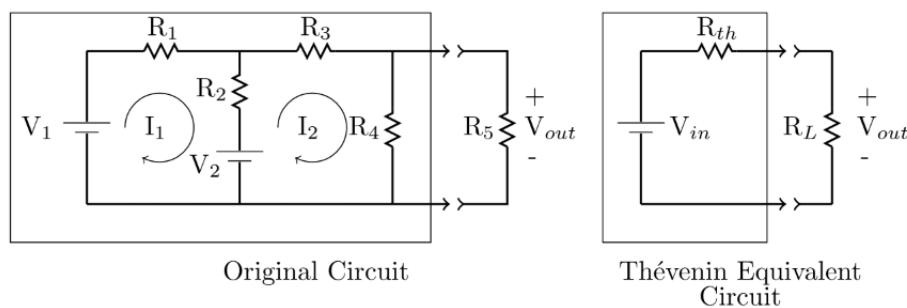


Figure 7.9: Thévenin equivalent circuit.

There is a standard procedure for calculating V_{th} and R_{th} . The Thévenin equivalent voltage is equal to the voltage that would be found across the terminals without anything attached to it. Here, R_5 is the reference component so it is placed with an open circuit. So the voltage across R_4 is the voltage being output by this circuit and is equal to V_{th} . In this example the Thévenin voltage is equal to the product of the current in the second loop, I_2 , and the resistance R_4 . If the

value for all the voltage sources and resistors are known, the current across R_4 can be found by solving for the loop currents. This will yield the Thévenin voltage as being

$$V_{th} = I_2 R_4 \quad (7.1.11)$$

To find R_{th} , R_5 is again removed and all voltage sources in the circuit are replaced with their internal resistances. In this case the voltage sources are assumed to be the ideal and are replaced with short circuits. The Thévenin resistance is then the equivalent resistance at the open terminals. For the circuit in Figure 7.9 this is

$$R_{th} = ((R_1 \parallel R_2) + R_3) \parallel R_4 \quad (7.1.12)$$

where \parallel indicates the resistors are to be added in parallel.

Another way to measure V_{th} and R_{th} is by knowing that the Thévenin equivalent circuit in Figure 7.9 is identical to the circuit depicted in Figure 7.8. Equation 7.1.10 can therefore be used to find the Thévenin voltage and resistance, V_{th} and R_{th} , by varying the load resistance, R_5 . Here, V_{out} is the voltage across R_5 , and V_{in} and R_i are V_{th} and R_{th} respectively. Again, plotting $\frac{1}{V_{out}}$ versus $\frac{1}{R_5}$ should yield a straight line. This method produces a graph giving R_{th} and V_{th} from the slope and intercept.

So by attaching a load resistor to a circuit, in the case of Figure 7.9 this was R_5 , the voltage across the resistor is dependent on the internal resistance as explained by the voltage divider rule. The effect that a load resistor has on circuits is called **circuit loading**. Imagine the basic voltage divider in Figure 7.5 where a voltmeter is used to measure V_{out} . Just like the power supply, voltmeters have some internal resistance. In the analysis of Equation 7.1.5, it was assumed that no current goes into the voltmeter, which is only true for an ideal voltmeter. The effect that non-ideal voltmeters have on voltage dividers is shown in Figure 7.10

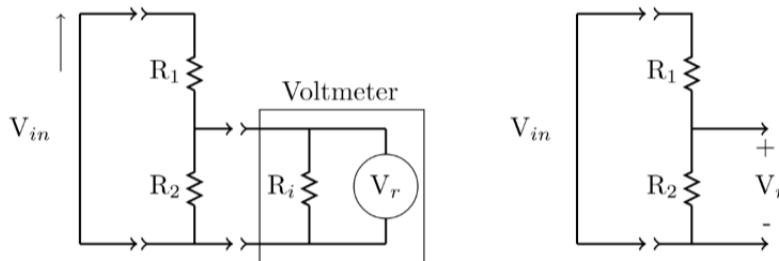


Figure 7.10: Internal resistance of voltmeter.

This schematic is similar to Figure 7.4 with the exception that the voltmeter now has an internal resistance R_i . Ideally, the voltage across R_2 is the ratio of R_2 to the entire resistance $R_1 + R_2$. Once a voltmeter is connected, the voltage is split between R_1 and the combined resistance of R_2 and R_i , in parallel. In Figure 7.10, V_r is the voltage read by the voltmeter and V_{out} , although not explicitly shown, is the actual voltage across R_2 without the voltmeter attached. This equivalent resistance is given by

$$R_{eq} = R_2 \parallel R_i = \frac{R_2 R_i}{R_2 + R_i} \quad (7.1.13)$$

It is seen that if R_i is much larger than R_2 , then R_{eq} is nearly equal to R_2 , and Equation 7.1.5 still holds. This is what would be expected from an ideal voltmeter. If this is not the case then R_{eq} will be smaller than R_2 . So a larger portion of voltage will develop across R_1 than before the voltmeter was connected. The voltage measured by the voltmeter will now be smaller than expected. This effect is called ***meter loading*** and is seen when R_2 in Equation 7.1.5 is replaced with R_{eq} to get

$$V_r = V_{in} \frac{R_{eq}}{R_{eq} + R_1} \neq V_{out} \quad (7.1.14)$$

Equation 7.1.14 gives the output voltage, taking the effect of the voltmeter on the circuit into account. It can be used to calculate the internal resistance of the voltmeter. Furthermore, it is possible to rearrange Equation 7.1.14 to find the voltmeter internal resistance by

$$R_i = \frac{R_1 R_2}{R_2 \left(\frac{V_{in}}{V_r} - 1 \right) - R_1} \quad (7.1.15)$$

If R_1 , R_2 and V_{in} are known, and V_r is the reading from the voltmeter.

Two multimeters are used in this experiment, the Philips digital meter and the Sanwa analog meter. The Philips multimeter is close to ideal in many situations and is used for most of the measurements in this experiment. The Sanwa multimeter, which is farther from ideal, is used to see the effects of meter loading.

In this experiment the Philips multimeter is used for taking voltage and resistance measurements. All connections are made to the red V jack and the black COM jack. Switching between voltage and resistance readings is done via buttons along the bottom of the multimeter, under the display. The DC voltage function is selected by the V button and the resistance function is chosen with the 2W button. The three top right buttons on the instrument are ranging controls. The range can be changed to automatic or manual with the AUT/MAN button. In manual mode the UP and DOWN buttons are used to change the range. Auto ranging is recommended unless otherwise specified. When measuring resistance be sure to disconnect any power supplies from the circuit. The internal resistance of the Philips voltmeter is $10 \text{ M}\Omega$.

The Sanwa multimeter is an analog meter with a fairly low input resistance. This multimeter will be used to witness the effect of meter loading. The internal resistance of the Sanwa multimeter is directly proportional to the voltage range it is reading in and is $20 \text{ k}\Omega/\text{V}$.

7.2.0 Experimental Procedure

Step 1: Before constructing any circuits measure the resistance of all resistors using the Philips multimeter.

- Construct the basic voltage divider depicted in Figure 7.11 with $R_D = 100 \text{ }\Omega$ and $R_1 = 20 \text{ }\Omega$. Turn on the power supply and connect the Philips multimeter across R_1 and take measurements of V_{out} while changing V_{in} from 1.0 V to 5.0 V .

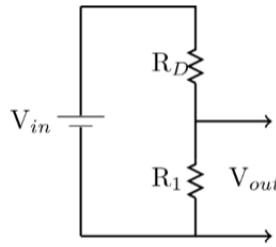


Figure 7.11: Schematics of the basic voltage divider

Step 3: Measure the resistance of the potentiometer R_x using the Ohmmeter across the red and black terminals. Build the circuit in Figure 7.12 with $R_D = 100 \Omega$ and $R_1 = 20 \Omega$. Set V_{in} to approximately 1.0 V and measure V_{in} with the Phillips multimeter.

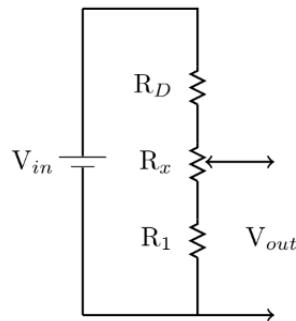


Figure 7.12: Voltage divider with three resistors

Step 4: Connect $R_D = 100 \Omega$ to the positive output terminal of the power supply. This will show the effects of having a power supply with a high output resistance. Connect a resistance decade box in series to serve as a variable load. The setup should resemble Figure 7.8. Set the power supply to 1.0 V. Using the Philips multimeter, take at least 10 measurements of V_{out} while varying the load resistance from 10Ω to $10 k\Omega$.

Step 5: Remove R_D from the circuit. Now R_i is the actual internal resistance of the power supply to be measured. Repeat the measurements from Step 4.

Step 6: Choose one of the circuits in Figures 7.13-7.14 and build it. Set the power supply between 1.0 V and 5.0 V. Measure V_{th} directly with a voltmeter. Replace the power supply with a short and measure R_{th} directly with an ohmmeter. **Remember to disconnect the power from a circuit before taking measurements of resistance with an ohmmeter.**

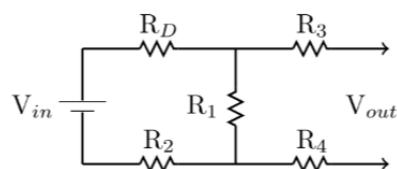


Figure 7.13: Optional circuit diagram A

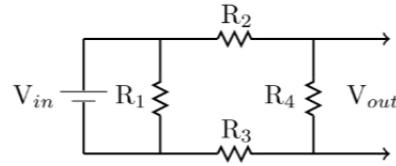


Figure 7.14: Optional circuit diagram B

Step 7: Connect a resistance decade box across V_{out} for the circuit chosen. This is the load resistor R_L and creates a voltage divider between R_{th} and R_L . Take at least 10 measurements of V_{out} across the load resistor, as the resistance is varied from 10Ω to $10 k\Omega$. Equation 7.1.10 can then be used to deduce the Thévenin equivalent voltage and resistance.

Step 8: Using resistance measurements only, deduce the contents of the box pictured in Figure 7.15.

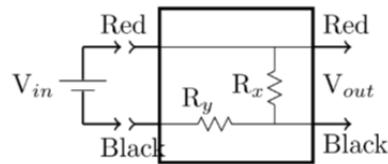


Figure 7.15: Voltage divider black box

- Using the same black box as Step 8, connect the positive end of the power supply to the red terminal, and the negative end to the black terminal. Set V_{in} to 3 V and take voltage readings across the red and white terminals. Call this voltage V_r . Measure V_r with the Philips multimeter at voltage ranges 3 V, 30 V and 300 V

Step 10: Measure V_r with the Sanwa analog meter at voltage ranges 0.5 V, 2.5 V, 10 V and 50 V

7.3.0 Error Analysis

The wires used are assumed to be ideal, but they do have a small amount of resistance. For some steps such as finding the internal resistance of the power supply, the resistance in the wires may affect the value obtained. Another source of uncertainty arises from the self heating of the resistors when current passes through them. Also for a 5 digit multimeter like the Philips instrument, external noise can be seen in the variance of the lower digits. In this case, the error can be estimated as half the smallest non-varying digit. The uncertainties for voltage and resistance measurements for the multimeter are given in Tables 7.1 and 7.2, respectively. The Sanwa multimeter accuracy is given by half the resolution of the scale for that range.

Table 7.1: The accuracy of the voltage measurements for the Philips multimeter

Range	% of reading	% of range
300 mV	0.0025	0.0013
3 V	0.0020	0.0010
30 V	0.0025	0.0013
300 V	0.0025	0.0010

Table 7.2: The accuracy of the resistance measurements for the Philips multimeter

Range	% of reading	% of range
3 kΩ	0.01	0.0033
300 kΩ	0.01	0.0033
3 MΩ	0.02	0.0033

Chapter 8

Fourier Series

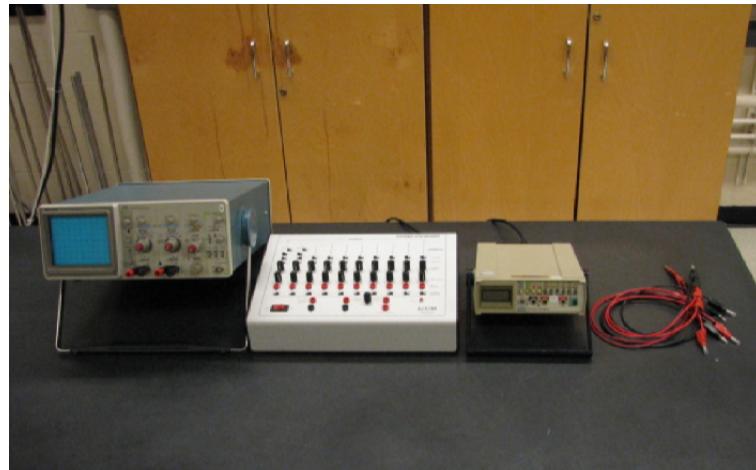


Figure 8.1: Equipment used for Fourier series experiment

8.1.0 Background

All sounds are pressure waves transmitted through the air. One method of describing these waves is to make use of a ***Fourier series***. These series are composed of a multitude of sine and cosine functions of differing frequencies and amplitudes. Fourier series are used to deal with problems in such diverse fields as acoustics, geophysics, and astronomy. In 1747 D'Alembert formulated the general equation of a vibrating string. In 1753, Daniel Bernoulli formulated a solution of the vibrating string problem based on a series of sines and cosines. This started a debate whether smooth waves could be added up to produce waves with sharp jumps and corners. At first the smoothness problem seemed minor, but this same series began to appear elsewhere. In 1807, Fourier submitted a paper that explained how to use this series to solve problems in heat transfer. This paper was criticized, because researchers did not accept that smooth waves could sum to jagged waves. Later analysts proved that a series sum of smooth functions could indeed produce a non-smooth result. Today, these series of sines and cosines are used routinely and are named ***Fourier series*** in honour of Jean-Baptiste Fourier.

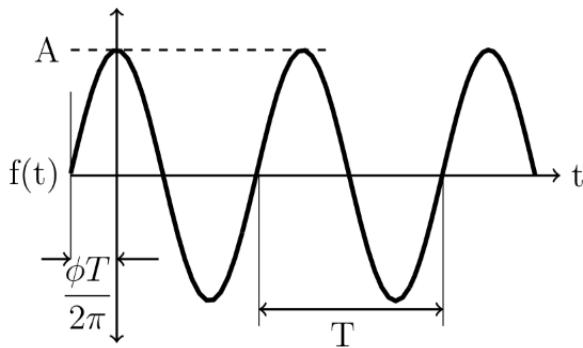


Figure 8.2: Phase shifting of a sine wave.

Fourier series are based on the sine function. Sine is an example of a periodic function. It repeats at a regular interval called a period. As can be seen in Figure 8.2, the period, T , is the length of time for one cycle of the wave. The frequency, $f = 1/T$, is the number of cycles in one second. A sine wave has two other parameters that can be varied. The amplitude, A , determines the height of the wave. The phase shift, ϕ , is the amount the wave is shifted left or right. Sine waves are depicted mathematically by:

$$f(t) = A \sin\left(\frac{2\pi}{T}t + \phi\right) \quad (8.1.1)$$

The quantity $2\pi/T = 2\pi f = \omega$ is called the angular frequency.

Sine waves are added together to form Fourier series. In this light, the addition properties of two sine waves is a special case of Fourier series. If the frequency and the phase are identical, then the amplitude of the sum is the sum of the amplitudes. The frequency and phase remain unchanged. This means that

$$A \sin\left(\frac{2\pi}{T}t + \phi\right) + B \sin\left(\frac{2\pi}{T}t + \phi\right) = (A + B) \sin\left(\frac{2\pi}{T}t + \phi\right) \quad (8.1.2)$$

The addition of sines with the same frequency but different phases introduces new properties. Suppose the two signals are separated by a phase of $\pi/2$. Then

$$A \sin\left(\frac{2\pi}{T}t\right) + B \sin\left(\frac{2\pi}{T}t + \frac{\pi}{2}\right) = C \sin\left(\frac{2\pi}{T}t + \phi\right) \quad (8.1.3)$$

where

$$C = \sqrt{A^2 + B^2} \quad (8.1.4)$$

and

$$\phi = \tan^{-1}\left(\frac{B}{A}\right) \quad (8.1.5)$$

The amplitude and phase of the sum are now functions of the amplitude and phase of the addends and are given by Equations 8.1.4 and 8.1.5. However, the frequency remains unchanged and the sum is still a sinusoidal wave.

When sinusoidal waves of different frequencies are added, the sum is no longer a sine wave. For example, if sines of identical amplitude and unequal but close frequency are added, the

result takes the form shown in Figure 8.3. This is the phenomenon known as **beats**. A lower frequency envelope is seen to be riding on the outside of a higher frequency signal, known as a **carrier**, because it “carries” the lower frequency beat signal along.

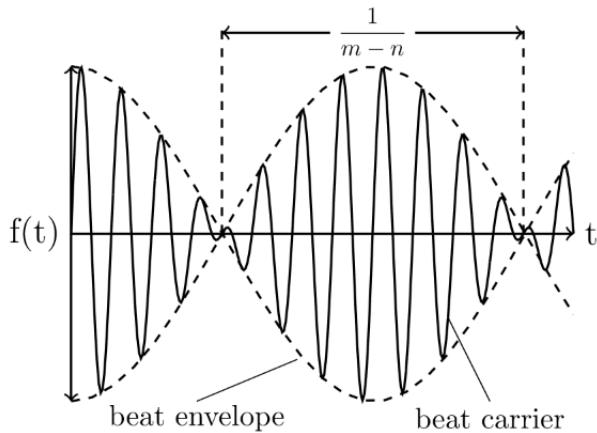


Figure 8.3: An illustration of beat frequency

Mathematically, adding a sine wave of frequency n to a sine wave of frequency m yields the expression:

$$\sin(nx) + \sin(mx) = 2 \sin\left[2\pi\left(\frac{m+n}{2}\right)x\right] \cos\left[2\pi\left(\frac{m-n}{2}\right)x\right] \quad (8.1.6)$$

If m and n are close together, then Equation 8.1.6 can be thought of as a time varying amplitude applied to a sine wave of frequency $(m+n)/2$. Although this time varying amplitude has frequency $(m-n)/2$, Figure 8.3 shows that the frequency at which the minimum value appears is twice this amount. So the beat envelope is headed as the $m - n$.

The results of Equations 8.1.2 through 8.1.6 are found by use of standard trigonometric identities and results in waves that are repetitive in nature. However, the resulting sum of two sine waves need not repeat itself. An interesting fact about the addition of different frequencies, m and n , is that the resulting wave is periodic only if there exists two rational numbers p and q such that

$$pm = qn \quad (8.1.7)$$

i.e. if the numbers m and n are commensurable. If the frequencies m and n are commensurable, then the frequency of the sum is given by the **greatest common divisor** of m and n .

A Fourier series generalizes the sum of two sine waves to the sum of an arbitrary number of sine and cosine waves. Each element of the series can have a difference amplitude. However, the frequency of any element must be an integer multiple of one frequency known as the **fundamental**. The integer multiples of the fundamental are known as **harmonics**. Suppose the fundamental has a period T . Then a Fourier series has the form

$$f(t) = a_0 + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{2\pi n}{T}t\right) + b_n \sin\left(\frac{2\pi n}{T}t\right) \right) \quad (8.1.8)$$

The constants $a_0, a_1, a_2, \dots, b_1, b_2, \dots$ are called **Fourier coefficients**. The most significant property of Fourier series is that any function, $f(t)$, that repeats with a period T can be represented

by such a series (assuming it doesn't have infinite discontinuities). This includes waves with sudden jumps and sharp corners.

A crucial problem is how to find the Fourier coefficients a_0, a_n and b_n for a given function $f(t)$, $n = 1, 2, 3, \dots$. The Fourier coefficients are calculated by the formulae

$$a_0 = \frac{1}{T} \int_0^T f(t) dt \quad (8.1.9)$$

$$a_n = \frac{2}{T} \int_0^T f(t) \cos\left(\frac{2\pi n}{T}t\right) dt \quad (8.1.10)$$

and

$$b_n = \frac{2}{T} \int_0^T f(t) \sin\left(\frac{2\pi n}{T}t\right) dt \quad (8.1.11)$$

It can be seen from Equations 8.1.9, 8.1.10, and 8.1.11 that the Fourier coefficients for a function $f(t)$ are found by some form of integration over the period T . The reason for this is that certain integrals have the useful property of being zero for all terms of the Fourier series except one. They act as a filter that preserves one term of the Fourier series and eliminates all the others. In particular, we use the orthogonality relation for sine and cosine

$$\int_0^T \cos\left(\frac{2\pi n}{T}t\right) \sin\left(\frac{2\pi m}{T}t\right) dt = 0 \quad \forall m, n \in \mathbb{N} \quad (8.1.12)$$

$$\int_0^T \cos\left(\frac{2\pi n}{T}t\right) \cos\left(\frac{2\pi m}{T}t\right) dt = \frac{L}{2} \delta_{mn} \quad (8.1.13)$$

and

$$\int_0^T \sin\left(\frac{2\pi n}{T}t\right) \sin\left(\frac{2\pi m}{T}t\right) dt = \frac{L}{2} \delta_{mn} \quad (8.1.14)$$

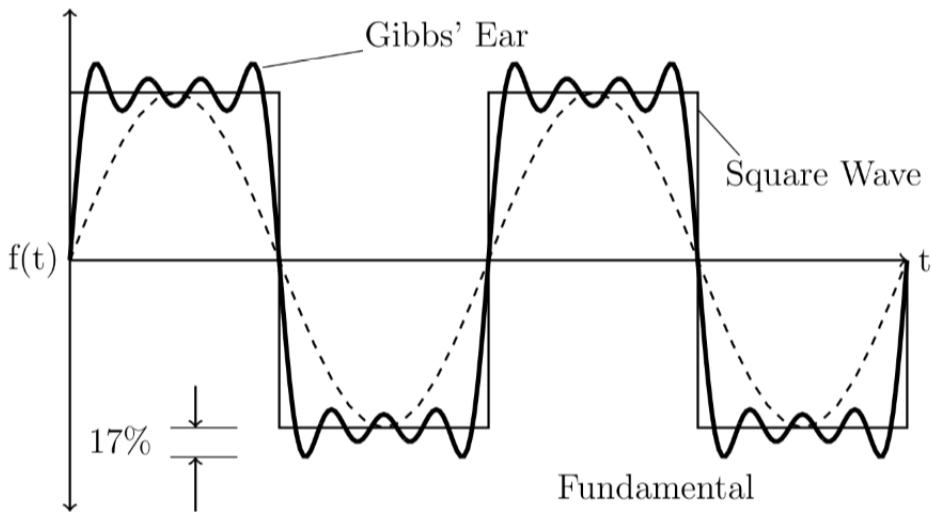


Figure 8.4: An illustration of Gibbs' phenomenon.

In Figure 8.4 it can be seen that the shape is converging to a square wave. However, there seems to be a greater difficulty for the Fourier series to track at the discontinuity. There is a

sudden jump or overshoot just before each rise and fall. This propensity for a Fourier series to overshoot at the edges is called **Gibbs' phenomenon** or **Gibbs Ears**. It turns out that this behaviour is a general property of all Fourier series. In fact, as more terms are added the ears get narrower, but their height may not decrease. For the square wave, it can be shown that the height of the ears never falls below 17% above the square wave level.

This experiment examines the wave phenomena described here using an instrument called a **Fourier Synthesizer**. A Fourier synthesizer is an instrument that creates arbitrarily shaped periodic signals by use of Fourier series. The device generates two sine waves at a fundamental frequency of 440 Hz, as well as eight other harmonic sine waves at frequencies 880 Hz, 1320 Hz, ... up to 3960 Hz. The phase and amplitude of each of these ten signals can be independently adjusted. Any combination of them can be added together. In this way the synthesizer can generate the first nine terms of any Fourier series.

The Fourier synthesizer used in this experiment is shown in Figure 8.5. There are ten columns of controls. One column for each of the two fundamentals and eight harmonics. At the top of each column are the phase controls. Two switches select 90° and 180° phase shifts. The third phase control is continuously variable over a 90° range. Together, these three controls provide a full 360° phase adjustment range. Below the phase controls is an amplitude control to set the size of the signal. An in/out switch selects whether the column gets added to the output or not. Each column also includes an output point where the signal generated by that column can be measured.

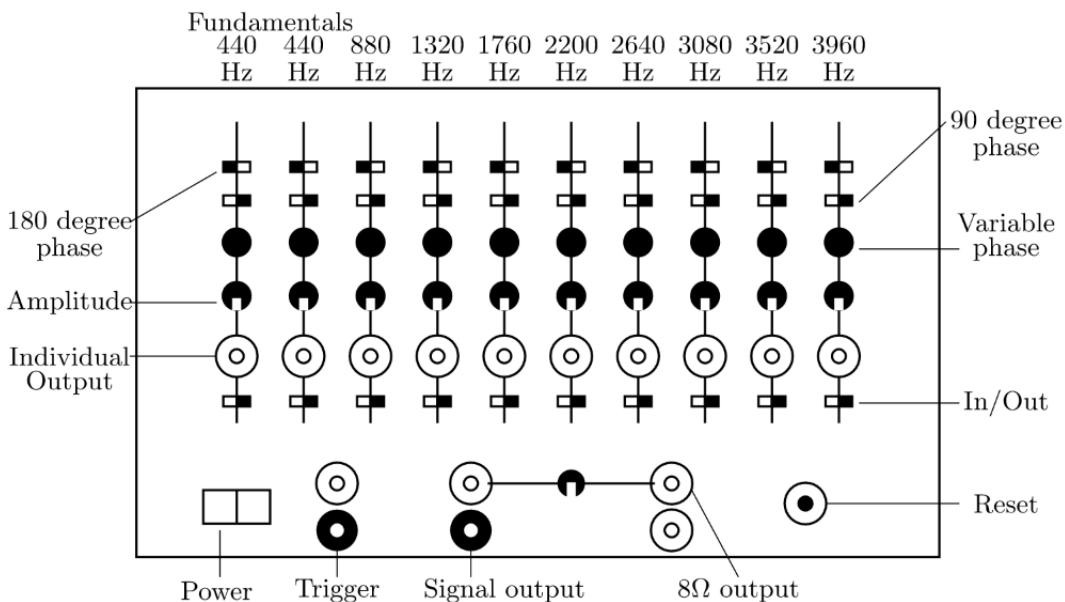


Figure 8.5: A schematic picture of the Fourier synthesizer.

The third output is called a **trigger**. This signal is a square wave locked in phase with the fundamental. The edges of this square wave represent a phase angle of 0°. This signal serves as a time marker (metronome) against which phases can be measured. Lastly, the reset button is used to periodically resynchronize all the harmonics back together.

The synthesizer can be used to experimentally examine the properties of the addition of sine

waves. Equation 8.1.3 can be examined by adding a sine wave to another phase shifted sine wave of the identical frequency. The amplitude and phase of the sum would be expected to vary in the systematic manner suggested by Equations 8.1.4 and 8.1.5. Equation 8.1.6 and Figure 8.3 can be tested by observing the resulting waveform when sines of identical amplitude but differing frequency are added together. For best results, the chosen frequencies should be relatively close together. This suggests adding together high harmonics such as 5, 6, 7, 8, and 9. The resulting frequencies of the beat envelope and the beat carrier can be measured with an oscilloscope and compared against Equation 8.1.6.

Similarly, Equation 8.1.7 can be examined by adding together two sinusoidal functions of arbitrary amplitude, phase, and frequency. The sum would be expected to have a frequency given by the greatest common divisor, independent of the amplitude and phase.

Most importantly, Equations 8.1.9, 8.1.10, and 8.1.11 can be used to find Fourier coefficients for various signals. It should be noted that the resulting shape is quite sensitive to the actual amplitude and phase of each harmonic. For best results, the amplitude of each harmonic should be set using the multimeter and the phase should be set using the oscilloscope.

8.2.0 Experimental Procedure

Step 1: Connect the Fourier synthesizer to the oscilloscope. One channel on the oscilloscope is best left connected to the trigger. The second channel on the oscilloscope is used to view the signal output or the individual output of each harmonic, whichever is most useful. The multimeter is connected to the individual output of whichever wave need to be measured at the time. A common ground should be used between all the equipment. For best results, the oscilloscope should be AC coupled.

Step 2: To test Equation 8.1.3, choose one fundamental to have an amplitude A and the other amplitude B . Set the phase of A to 0° and the phase of B to be 90° . The phase is set using the oscilloscope, by observing the alignment of the sine wave against one edge of the trigger. A phase shift of 90° is $1/4$ of a period. Since the period is $1/440$ s, the time shift for 90° is $1/4 \cdot 440$ s ≈ 570 μ s. Fix the amplitude A to some known value using the multimeter. Then, for at least eight different amplitudes of B , measure the amplitude of C and the phase of the sum. Amplitude is best measured with the AC voltmeter, and phase is best measured with the oscilloscope.

Step 3: Choose any one of the following tests for the addition of sine waves of different frequency:

- (A) Examine beats as given by Equation 8.1.6 and Figure 8.3. Choose any two neighboring higher harmonics and set their amplitudes to be equal and the phases to be 0° . Measure the frequency of the beat envelope and beat carrier. Repeat for at least two other neighboring harmonics.
- (B) Examine the sum frequency predicted by Equation 8.1.7. Choose any two frequencies with arbitrary non zero amplitudes and arbitrary phase. Measure the period of their sum. The sum frequency can be compared to the expected value given by the greatest common divisor of the individual frequencies. Repeat for at least two other combinations.

Step 4: The synthesizer can generate a 9 term Fourier series for the square wave, as seen in Figure 8.4. A recipe for synthesizing the square is given in Table 8.1. To synthesize a particular wave shape, each harmonic 1 through 9 must be set to a particular amplitude and phase. Each amplitude in the table is a percentage relative to a maximum amplitude within the first nine terms. Each phase angle is the number of degrees by which to shift that harmonic. Again, for best results, the AC voltmeter should be used to set the amplitude and the oscilloscope used to set the phase.

Table 8.1: Amplitudes and phase shifts required to synthesize a square wave.

n	1	2	3	4	5	6	7	8	9
% Ampl.	100	0.0	33.3	0.0	20.0	0.0	14.3	0.0	11.1
Phase	90°	0°	90°	0°	90°	0°	90°	0°	90°

Step 5: Gibbs ears can be measured once the square wave has been synthesized. As each harmonic above number 1 is “switched in” record the height and width of Gibbs’ ears. This data can be used to check whether the width steadily decreases, while the height goes no lower than 17%.

Step 6: Synthesize one of the following waveforms. When synthesizing your signal do not trust the amplitude dials on the synthesizer. Instead set the amplitudes individually by measuring them with a multimeter.

Table 8.2: Amplitudes and phase shifts required to synthesize a ramp.

n	1	2	3	4	5	6	7	8	9
% Ampl.	100	78.5	11.1	39.3	4.0	26.2	2.0	19.6	1.2
Phase	0°	270°	0°	270°	0°	270°	0°	270°	0°

Table 8.3: Amplitudes and phase shifts required to synthesize a triangle.

n	1	2	3	4	5	6	7	8	9
% Ampl.	100	0.0	11.1	0.0	4.0	0.0	2.0	0.0	1.2
Phase	0°	0°	0°	0°	0°	0°	0°	0°	0°

Table 8.4: Amplitudes and phase shifts required to synthesize a sawtooth.

n	1	2	3	4	5	6	7	8	9
% Ampl.	100	50	33.3	25.0	20.0	16.7	14.3	12.5	11.1
Phase	90°	90°	90°	90°	90°	90°	90°	90°	90°

Step 7: Synthesize one of the following rectified sine waves.

Table 8.5: Amplitudes and phase shifts required to synthesize a full rectified sine wave.

n	1	2	3	4	5	6	7	8	9
% Ampl.	0	100	0	20.0	0.0	8.6	0	4.8	0
Phase	90°	0°	90°	0°	90°	0°	90°	0°	90°

Step 8: Three methods of transmitting binary signals (ones and zeros) are frequency shift keying (FSK), amplitude shift keying (ASK), and phase shift keying, as shown in Figures 8.6, 8.7, and 8.8. If ASK is used, one cycle of a sine wave represents a one and a flat line is a zero. Likewise, FSK represents a one by one cycle of a sine wave and represents a zero by one cycle at a different frequency sine wave. For PSK, a one is designated by one cycle of a zero phase and a zero is one cycle of a 180° phase shifted sine wave.

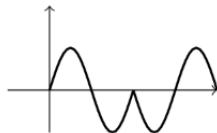


Figure 8.6: Pulse shift keying

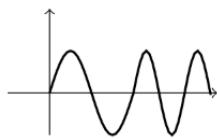


Figure 8.7: Frequency shift keying

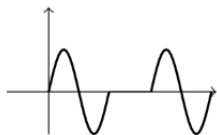


Figure 8.8: Amplitude shift keying

Table 8.6: Amplitudes and phase shifts required for PSK.

n	1	2	3	4	5	6	7	8	9
% Ampl.	100	0.0	60.0	0.0	14.3	0.0	6.7	0.0	3.9
Phase	0°	0°	180°	0°	180°	0°	180°	0°	180°

Table 8.7: Amplitudes and phase shifts required for FSK.

n	1	2	3	4	5	6	7	8	9
% Ampl.	56.8	100.0	41.6	22.0	8.4	0.0	2.9	2.1	0.0
Phase	30°	150°	90°	30°	330°	0°	30°	330°	0°

Table 8.8: Amplitudes and phase shifts required for ASK.

n	1	2	3	4	5	6	7	8	9
% Ampl.	84.9	100.0	50.9	0.0	12.1	0.0	5.7	0.0	3.3
Phase	0°	90°	180°	0°	180°	0°	180°	0°	180°

8.3.0 Error Analysis

Generally, the errors in an experiment arise from systematic errors in the apparatus and measurement errors in the instruments. In this experiment, the measurement errors are determined by the accuracy of the oscilloscope and the multimeter. The systematic errors arise from the behaviour and performance of the Fourier synthesizer. The measurement errors of the oscilloscope and multimeter can be readily estimated.

However, there are many possible sources of systematic error lurking inside the synthesizer. The fundamental and harmonic sine waves generated by the synthesizer have some distortion. Also, the summation of all the components is not mathematically perfect. Furthermore, the synchronization of the harmonics to each other and to the fundamental is not exact. The magnitude of these and other internal errors is essentially unknown.

Chapter 9

AC Circuit and RLC Resonance

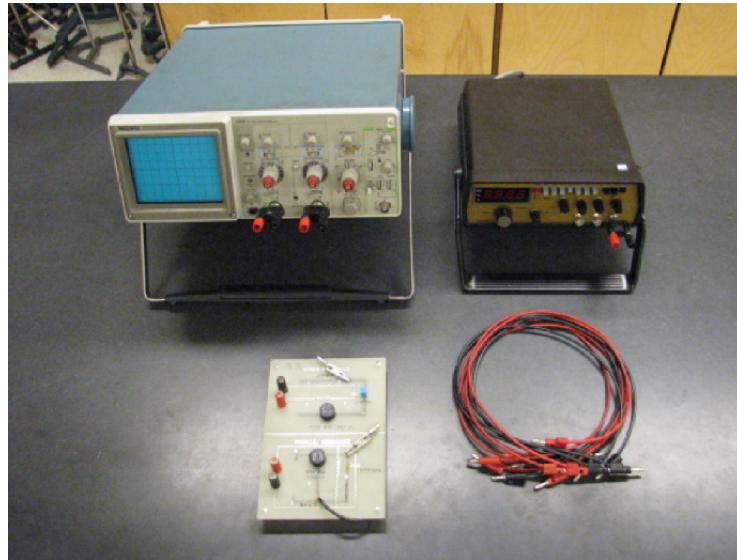


Figure 9.1: Equipment used to investigate a resonance in RLC circuits.

9.1.0 Background

The current delivered to homes and factories is an oscillating function of time. This is called **alternating current**, or **AC**. All the appliances connected to ordinary outlets in homes therefore involve circuits with oscillating currents. Furthermore, electronic devices, such as radio transmitters and receivers, involve a variety of circuits with oscillating currents of high frequency. Many of these circuits have natural frequencies of oscillation. Such circuits exhibit the phenomenon of **resonance** when the natural frequency matches the frequency of a signal applied to the circuit. For instance, the tuning of a radio relies on an oscillating circuit whose frequency of oscillation is adjusted by means of a variable capacitor so that it matches the frequency of the radio signal.

A common circuit known as the **RLC circuit** exhibits the phenomenon of resonance. It is named RLC because it includes a resistor, a capacitor, and an inductor. In this laboratory we will only analyse a series RLC circuit.

9.1.1 Series RLC Circuits

The series RLC circuit used in this experiment is shown in Figure 9.2. The values of resistance, R_r , capacitance, C , and inductance, L , are written on the components themselves. The total resistance of the circuit, R , is the sum of the resistance of the resistor, R_r , and the inductor (solenoid), R_L .

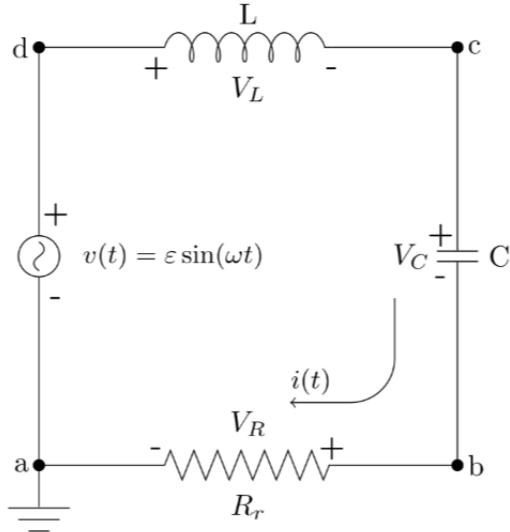


Figure 9.2: Series RLC circuit.

Assume that the internal resistance of the function generator can be ignored and choose the input signal to be $v(t) = \varepsilon \sin(\omega t)$, where ε is the maximum emf. The emf acts as a driving force that moves the charge in the circuit. An application of Kirchhoff's voltage law gives

$$\sum V = 0 = V_L + V_C + V_R - \varepsilon \sin(\omega t) \quad (9.1.1)$$

where

$$V_C = \frac{1}{C} \int i \cdot dt \quad (9.1.2)$$

$$V_L = L \frac{di}{dt} \quad (9.1.3)$$

and

$$V_R = iR \quad (9.1.4)$$

are the AC voltages across each component.

Substituting Equations 9.1.2, 9.1.3, and 9.1.4 into Equation 9.1.1 and rearranging gives the expression

$$\varepsilon \sin(\omega t) = \frac{1}{C} \int i \cdot dt + L \frac{di}{dt} + iR \quad (9.1.5)$$

Differentiating both sides with respect to time, t , to remove the integral, yields the differential equation

$$\varepsilon \omega \cos(\omega t) = \frac{i}{C} + L \frac{d^2i}{dt^2} + R \frac{di}{dt} \quad (9.1.6)$$

The steady state solution of this equation is

$$i(t) = \frac{\epsilon}{\sqrt{R^2 + (\omega L - \frac{1}{\omega C})^2}} \cdot \sin(\omega t - \phi) \quad (9.1.7)$$

where

$$\phi = \tan^{-1} \left[\frac{1}{R} \left(\omega L - \frac{1}{\omega C} \right) \right] \quad (9.1.8)$$

The same equation can be obtained through the analysis of the circuit's complex impedance.

The **angular frequency**, ω , is $2\pi f$, where f is the frequency in cycles per second, or Hertz (Hz). Equations 9.1.7 and 9.1.8 imply that the current flowing in a series RLC circuit is a sine wave with an amplitude that changes with frequency and a phase difference from the signal generator that also changes with frequency. So the behaviour of the circuit strongly depends on the frequency of the applied voltage.

For Equation 9.1.7, the current, $i(t)$, becomes an extremum when $\omega^2 = 1/LC$. The condition when the amplitude of $i(t)$ is an extremum is known as **amplitude resonance**. For series RLC circuits, the extremum turns out to be a **maximum**. Thus, at amplitude resonance the current flowing in the circuit has maximum amplitude. At all other frequencies, the current flowing in the circuit is smaller. This property makes the series RLC circuit useful in filters where a response at one frequency is preferred over all other frequencies. This is what happens in radio receivers when one station is selected. The frequency corresponding to the carrier of that station is chosen to be the amplitude resonant frequency.

When $\omega^2 = 1/LC$, Equation 9.1.8 implies that $\phi = 0$. The condition when $\phi = 0$ is called **phase resonance**. At phase resonance, the signals at points d and b in Figure 9.2 are in phase. So there is no phase difference between the applied voltage and the current through the circuit. In series RLC circuits, amplitude and phase resonance both happen at the **same frequency**. In other types of circuits this is not necessarily true.

The oscilloscope screen in Figure 9.3 shows two signals that are not in phase with each other. A phase difference between two waveforms is usually reported as an angle. Since horizontal distances on an oscilloscope are in terms of time, a conversion from time to angle has to be made. This is done by noting that the time for a full cycle is 2π rad, so the actual phase difference as an angle is found by determining it as a fraction of one cycle, and multiplying by 2π rad.

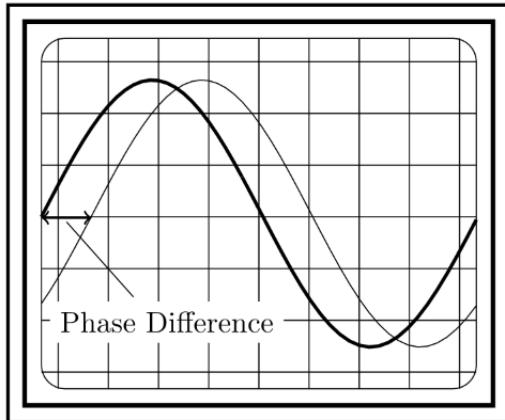


Figure 9.3: An illustration of phase difference

To measure a phase difference between two sine waves necessarily implies that they have the same frequency. The phase difference between two sine waves of various frequencies is not constant. Also, notice that the two sinusoidal waveforms depicted in Figure 9.3 have the same amplitude allowing for accurate measurement of the phase difference. Since in general, it is unlikely that any two signals have the same amplitude, a special control is available on the oscilloscope to shift amplitudes without requiring a change to the input signal. This control is called the vertical calibration adjustment. To make two signals have the same displayed amplitude simply modify the vertical calibration of one channel. After the phase measurement is completed, remember to return the vertical calibration setting back to normal, otherwise further vertical measurements will be incorrect.

When $i(t)$ becomes an extremum, the input frequency matches the natural frequency of the circuit. This is manifested as a voltage extremum on the oscilloscope screen. Thus, when searching for frequencies at which $i(t)$ becomes an extremum, look for frequencies corresponding to a voltage extremum on the oscilloscope screen. The reason for voltage measurements rather than current measurements is due to the fact that the oscilloscope only displays voltages. Moreover, the signal at point b is the voltage across the resistor, which is directly proportional to the current by Ohm's law.

9.2.0 Experimental Procedure

Step 1: Inspect the function generator and the dual channel oscilloscope. Make sure that the function generator is set to a sine wave output. Also, set the scope to show both channels on the screen simultaneously as this will be needed to verify the phase resonance of the circuits.

Step 2: For the series RLC circuit, setup the experiment as in Figure 9.4. Connect both the function generator and channel 1 of the oscilloscope between d and a , the input pins. Thus the trace on channel 1 will represent the input signal from the function generator. Connect channel 2 of the oscilloscope between b and a . The channel 2 trace will then be directly proportional to the current going through the circuit.

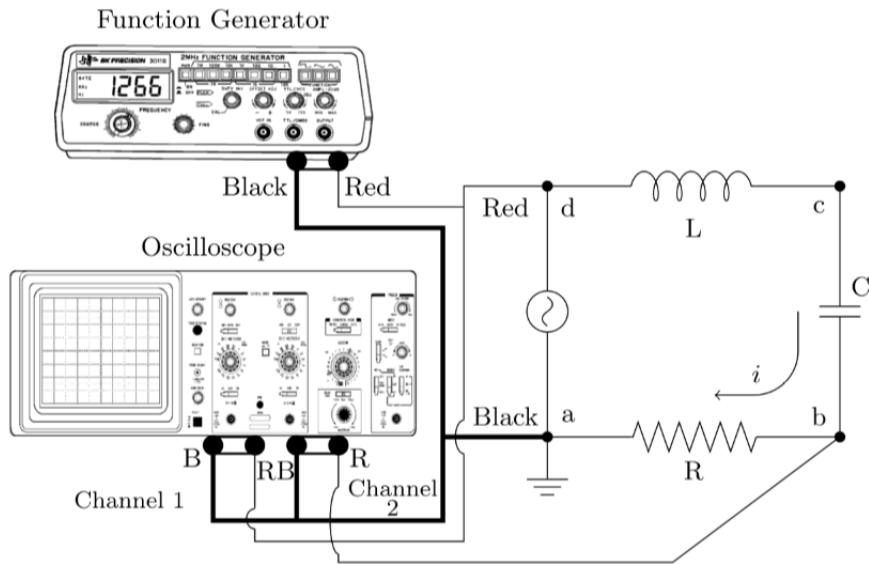


Figure 9.4: Series RLC circuit setup.

- Step 3: Find the amplitude/phase resonance by scanning through frequencies near the theoretical value. Stop varying the frequency when the channel 2 trace has its amplitude extremum. Record the experimental resonant frequency from the function generator. Verify that the frequency for the amplitude resonance is the same as the frequency for phase resonance ($\phi = 0$).
- Step 4: For the resonant frequency and at least 50 other frequencies, some below and some above the resonant frequency, record the peak-to-peak voltage value from the oscilloscope. Select your frequencies so you can precisely determine the maximum value and be able to observe the out-of-resonance behaviour of the circuit.

9.3.0 Error Analysis

The internal resistance of the function generator is ignored in this experiment. Consider all values written on the circuit boards as exact. The error in an oscilloscope measurement is half the resolution of the measurement scale. Remember to make detailed notes on which voltage scale and time scale the oscilloscope is set on when recording data.

There are two errors associated with the function generator frequency. A **drift error** caused by temperature variations inside the function generator and a **quantization error** which is associated with any digital instrument. The drift error can be estimated by observing variations in the output frequency. While you are making measurements, take note of the maximum and minimum frequency values displayed on the function generator. The difference between them divided by two gives the associated drift error. If the function generator has been on for some time then no drift in frequency values may be observable. In this case, the only error associated with the frequency will be the quantization error.

The error resulting from conversion of a frequency to a digital display is called the quantization error. Manufacturers usually give the value of this error as plus or minus one number in

the smallest digit of the display. The manual accompanying the B and K Precision Function Generator lists this value as ± 1 count.

Take the error in frequency to be the sum of the drift and quantization errors.

Chapter 10

AC Circuits and Voltage Dividers

10.1.0 Resistor-Capacitor Filters

There are two possible RC configurations, one with the output voltage taken over the resistor, and the other with the output taken across the capacitor. If the output is taken across the capacitor, we obtain a low pass filter, and if the output is taken across the resistor, we obtain a high pass filter.

Definition 10.1.1. *The frequency at which the magnitude of the ratio of output voltage to input voltage is $10^{-3/20} \approx 0.7079$ is called the cutoff frequency. This is the frequency at which half the input power is transferred to the output and is therefore sometimes called the -3 dB frequency.*

The cutoff frequency of a filter is considered to be the frequency at which the filter begins to block or attenuate the input signal.

Definition 10.1.2. *An all-pass filter requires two resistor-capacitor pairs whose products, (RC) , are equal. This configuration provides a constant ratio of V_{out} to V_{in} at all frequencies by rendering stray capacitance and inductance insignificant.*

Definition 10.1.3. *The gain function $G(\omega)$ is the transfer function of any circuit. It is the ratio of V_{out} to V_{in} (phase and amplitude count) as a function of frequency.*

10.2.0 Attenuator Probe Adjustments and Compensations

Since many measurements (high-impedance and high-frequency) require an attenuator probe, you should make sure it is compensated as a routine part of turning on the scope. A square wave intended for probe calibration and compensation is always available on the scope's front panel. Probe compensation is adjustable via a screwdriver slot on the probe body or connector.

The display of an undistorted square wave is a quick indication of correct compensation.

If we adjust the probe capacitor C_p to the correct value, a compensating amount of high frequency information will be bypassed around the probe resistor, and a square wave will have both flat tops and bottoms. We see spikes for too much high-frequency response from the probe, and curved edges for too little high frequency response.

Remark 10.2.1. In the context of high-frequency measurements, the total input impedance drops as the frequency increases because stray capacitance is in parallel to the probe. This is unavoidable and makes the $\times 10$ probe even more important in high-frequency environments to minimize loading errors.

Chapter 11

AC Circuits, Waveforms, and Phase

11.1.0 Measurement of Sine Wave Amplitudes

Digital multimeters (DMMs) are more convenient and accurate for measuring DC voltages, currents, and sinusoidal AC voltages and currents than oscilloscopes.

We can specify the amplitude of a sine wave in four ways: using the peak-to-peak value, the peak value, the average value, and the root mean square (RMS) value. The peak value is the value at the top of the waveform, the peak-to-peak is twice the peak, or the difference between the maximum and minimum. The RMS value is found by taking the square root of the average of the squared values of the current. This value of the sine wave current will produce the same heating effect in a resistor as a DC current of the same magnitude, since the power dissipated by a resistor is proportional to the square of the current running through it ($P = I^2R$).

The DMM measures the RMS voltage. The oscilloscope can be best used to measure the peak-to-peak voltage.

11.2.0 RC Circuit Current and Voltage Relationships

Remark 11.2.1. *Dual trace observation:* If there are two input signals, one to CH1 and one to CH2, you can only trigger on one signal or the other. To operate the oscilloscope such that you can observe two signals simultaneously you must select the vertical axis mode button [DUAL] and select either ALT or CHOP.

Remark 11.2.2. *Measuring phase difference:* Phase difference can be measured by the distance of separation between the signals when crossing the zero line. This time, divided by the time duration of one full cycle, defines the fraction of a cycle of phase difference. We know 1 cycle = 2π , so the fraction must be multiplied by the appropriate factor to be expressed in radians.

Chapter 12

Inductor-Resistor-Capacitor Filters

12.1.0 Quality Factor

The quality factor of an oscillator, Q , is a dimensionless parameter describing how slowly oscillations decay. More specifically, it roughly measures the number of oscillations (i.e. up to a constant) before the energy of the oscillator at a given time would decrease by a factor of e in the absence of a driving mechanism.

In the parallel resonant circuit, the Q value is directly proportional to the resistance. With a relatively large resistance, the circuit will be underdamped. This is characterized by exponentially decaying sinusoidal electrical oscillations in response to an impulse. Critical damping occurs with a slightly smaller resistance and is the point at which the circuit no longer oscillates, but returns in the minimum possible time to zero output after the impulse. Overdamping occurs with a small resistance, and is evidenced by capacitor charge returning more slowly than the case of critical damping.

When a parallel LRC circuit is used as a filter, the Q factor determines the bandwidth of the filter's response in frequency space. With a Q of 2 or 3, the response begins to be noticeably peaked. When Q is larger than 10, the filter's amplitude response is peaked by more than an order of magnitude at resonance. All measurements are traditionally normalized to the peak amplitude and to the frequency at which that peak occurs.

12.1.1 Important formulas

The impedance of an LRC circuit, with the components connected in parallel is

$$\frac{1}{Z_{LRC}} = \frac{1}{R} + \frac{1}{j\omega L} + j\omega C$$

and resonance frequency, f_0 , is given by

$$f_0 = \frac{1}{2\pi\sqrt{LC}}$$

the quality factor is

$$Q = \frac{f_0}{Bandwidth} = RC\omega_0$$

and the bandwidth is

$$Bandwidth = f_{high3dB} - f_{low3dB} = \frac{1}{2\pi RC}$$

Chapter 13

Transformer Voltage Gain, Ground Isolation

Chapter 14

Intro to Operational Amplifiers

Chapter 15

Inductor-Resistor-Capacitor Filters

Chapter 16

Bandwidth and Stability

16.1.0 Lecture Notes

Note that an ideal op-amp is a perfect linear differential operator. Additionally recall that $V_{out} = A(V_+ - V_-)$ at all frequencies. But if the open loop gain A is infinite always, as $\omega \rightarrow \infty$ we have $\frac{dV_{out}}{dt} \rightarrow \infty$. If there is any capacitance $q = CV$, $I \rightarrow \infty$ which is unphysical!

For any real op=amp we have $A = A(\omega)$ with $\lim_{\omega \rightarrow \infty} A(\omega) = 0$. We take a low pass filter on the gain: the open loop gain $A(\omega)$ is deliberately shaped as $1/\omega$ in the intended operating range.

In a log-log scale we have the following plot:

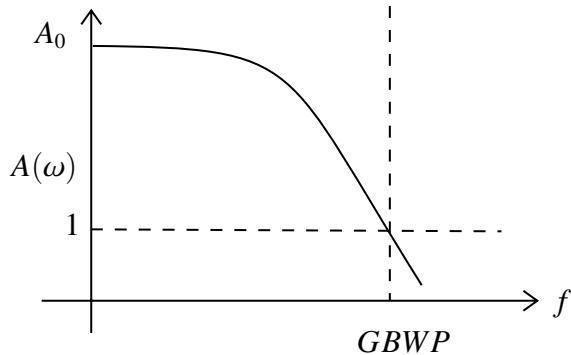


Figure 16.1: Open loop gain as a function of f in a log-log scale

The slope of the linear portion on this figure is $1/f$. The **Gain Bandwidth Product** (GBWP) is the frequency at which the gain $A(\omega)$ is equal to 1. We hope to express these ideas using some well known equations.

We consider the following op amp:

Recall $V_{out} = A(V_+ - V_-)$. We define $\beta = \frac{1}{Gain} = \frac{R_1}{R_1 + R_f}$. Then $V_{out} = A(V_{in} - \beta V_{out}) =$

$AV_{in} - A\beta V_{out}$), so $V_{out}(1 + A\beta) = AV_{in}$. so our closed loop gain is

$$\frac{V_{out}}{V_{in}} = \frac{A}{1 + A\beta}$$

Note 16.1.1. Not all op amps preserve the phase at the output. We need the negative feedback to be going back to the inverting input, or otherwise the devices will begin to oscillate.

16.1.1 Frequency Rolloff

The *frequency roll off* is generally $-20 \text{ dB}/\text{decade}$. In other words, the gain is down by a factor of 10 for every $10\times$ increase in frequency.

$$-20 \text{ dB} = 20 \log \left(\frac{V_{out}}{V_{in}} \right) \implies \frac{V_{out}}{V_{in}} = \frac{1}{10}$$

Equivalently this is -6 dB/octave , where an octave is a halving/doubling in frequency, so $\frac{V_{out}}{V_{in}} = 10^{-3/10}$.

Chapter 17

Basic OpAmp Difference Amplifier

17.1.0 Lecture Notes: Common Mode Rejection

We consider a difference amplifier.

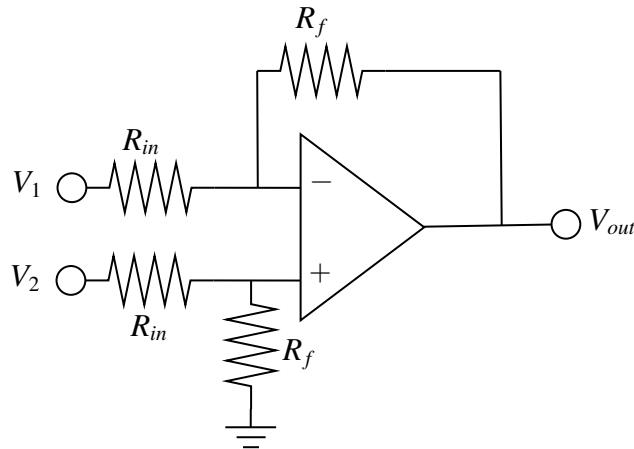


Figure 17.1: Difference Amplifier

We have the expression

$$V_{out} = \frac{R_f}{R_{in}} (V_2 - V_1)$$

Note for this to work we need to assume that our R_{in} 's and R_f 's are identical. Here the output is proportional to the difference in the two signal.

Now, the Common Mode Rejection Ratio (CMRR). In general we have $\frac{\text{Gain diff amp}}{\text{Gain common mode}} \sim 500$ to 50000. The gain of the diff amplifier is $\frac{V_{out}}{V_2 - V_1}$, while the common mode gain is $\frac{V'_{out}}{V_{in}}$ where here $V_2 = V_1 = V_{in}$. Both inputs have the same signal.

Chapter 18

Operational Amplifier Comparator

Chapter 19

Active Filters and Anti-Aliasing Filters

19.1.0 Lecture Notes: Sampling, Aliasing, and Analog Filters

Question 19.1.1. How can we accurately generate a set of discrete samples?

Recall the Nyquist theorem: to generate an unambiguous representation of a waveform at $\frac{1}{2}$ the sampling frequency we must limit the maximum frequency in the signal to NO GREATER than $\frac{1}{2}$ the sampling frequency.

We must avoid aliasing - where we have one frequency masquerading as another!

19.1.1 Filters

Question 19.1.2. How can we keep our maximum frequency within a safe frequency range?

When building a circuit we want about about a quarter of the Nyquist frequency gap between our frequencies after filtering and the Nyquist frequency. This is to ensure no frequencies appear above the Nyquist, f_N . A simple filter with a $\frac{1}{f}$ rolloff is not sufficient. We must construct a circuit of cascading filters to give a stronger rolloff.

Chapter 20

Swept Frequency Waveform Generator, DFT

Many laboratory function generators have the capability of changing frequency according to a voltage input. Most include an internal ramp generator to be used to smoothly tune the function generator across a known range of frequencies. This allows the automation of some otherwise tedious measurements.

Used alone, the swept frequency generator can simplify/automate the task of plotting the frequency response of filters. The oscilloscope can then be used to present an $X - Y$ plot of a simple RC or LRC filter's frequency response.

The Fourier Spectrum Scanner is an addition to the frequency generator. For the Fourier Spectrum Scanner, a multiplier is used in the “time domain” to produce sums and differences in the “frequency domain.” If we follow the multiplier with a low-pass filter, we are able to plot only the difference frequency. This can produce an $X - Y$ plot of the frequency components of complex time domain waveforms. By allowing the swept frequency function generator to automate the plotting of difference frequencies, we produce a Discrete Fourier Transform of the input waveform. The only requirement is that the input waveform be “stationary” during the duration of the scan.

20.1.0 Spectrum Analyzer

A signal multiplier will multiply the two input signals together. The result of the multiplication will be a superposition of signals that will have a frequency that is either the sum or the difference of the two input signals. For example, if the signal $f = \cos(\omega_1 t)$ is multiplied by the signal $g = \cos(\omega_2 t)$, the result will be

$$f * g = \cos(\omega_1 t) * \cos(\omega_2 t) = \frac{1}{2} [\cos((\omega_1 - \omega_2)t) + \cos((\omega_1 + \omega_2)t)]$$

The low-pass filter is designed to reject the signal with a summed frequency and to pass the signal with a difference frequency. Only the component of the signal with a difference frequency will be considered from here on. In this case $f * g \approx \frac{1}{2} \cos((\omega_1 - \omega_2)t)$. As the frequency of

the difference signal becomes small enough it starts to get passed by the low-pass filter. The gain response of the low-pass filter has a signal mound at $\Delta\omega = 0$, and goes to zero on either side. The maximum gain occurs at $\omega_1 = \omega_2$. The voltage out of the low-pass filter is a wave with a magnitude envelope that goes to zero away from $\Delta\omega = 0$. The frequency of the output signal decreases as the frequency of the difference signal approaches the frequency of the fixed signal.

Now consider the analysis of a square wave input signal. A square wave can be expressed as the sum of the odd harmonics of a sine wave

$$f_{square} \approx \frac{4}{\pi} \left[\sin(\omega_0 t) + \frac{1}{3} \sin(3\omega_0 t) + \frac{1}{5} \sin(5\omega_0 t) \right]$$

There will be a response across the low-pass filter for this signal every time the frequency of the swept signal gets close to the frequency of one of the harmonics.

Chapter 21

Lock-In Detector

Definition 21.0.1 (Noise). All disturbing elements over which the experimentalist has no control are noise. These elements include the fundamental thermal fluctuations of all matter not at absolute zero, statistical fluctuation due to the quantized nature of light and EM radiation, electrical currents, building vibrations, variations in room temperature, and stray electrical signals. These can be decreased to an arbitrarily small value but in practice are difficult to remove entirely from the picture.

Definition 21.0.2. The signal-to-noise ratio (S/N) in a system is determined by the signal and noise sources and by the bandwidths employed by each of the components of the system. The lock-in technique can improve the S/N in a system by 40 dB or more.

Lock-in detectors are used in cases where noise inhibits acceptable recovery of a signal. The primary role of the lock-in detector is the recovery of weak signals in the presence of noise. Essentially, it is a narrow band pass detector. The lock-in property of the lock-in detector functions as follows. The single frequency AC reference signal is used to modulate the source of the signal. This assures that the signal is related in frequency and phase to a reference voltage. Noise will be random with respect to the signal. The reference therefore maintains the centre of the phase-sensitive detector band-pass response at the signal frequency. Typical applications of the lock-in detector include measurements of weak light signals, measurements of conductance and capacitance in MOS structure, and nuclear magnetic resonance signal recovery. Lock-in detectors are commercially available.

Part II

Circuit Theory

Chapter 22

Introductory Examples

22.1.0 Direct Current and Circuit Laws

Current is the flow of charge. We will look at it as conventional current (positive flow). Voltage and current are constant for direct current (DC). Two various types of two-terminal circuits elements may be classified according to their terminal voltage-current $v(i)$ relationships:

- $v(i) = \text{constant}$
- $v = v(t), i = i(t)$ - ideal voltage source (independent voltage source)

Definition 22.1.1. The waveform $v(t)$ represents the voltage produced by the source. “Ideal source” means the voltage is maintained regardless of the drawn current.

Definition 22.1.2. An ideal battery has its voltage drop between the terminals being V_0 .



Voltmeter measures voltage (potential difference) between two points along the circuit.

Ammeters measure current through the circuit.

Definition 22.1.3 (Ohm's Law). Ohm's Law states that for a resistor connected between terminals A and B, the voltage drop from A to B is equal to the resistance multiplied by the current flowing from A to B through the resistor.

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

and

$$v(t) = Ri(t) \implies i(t) = \frac{1}{R}v(t)$$

for ohmic materials (such as wires)

22.1.1 Resistance

Resistance depends on the geometry and the material of the conductor. For a specific temperature, $R \propto \frac{L}{A}$, specifically

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

where ρ is the resistivity (Ωm).

Table 22.1: Material Resistivity

Material	Resistivity Ωm
Silver	1.59×10^{-8}
Copper	1.68×10^{-8}
Aluminum	2.65×10^{-8}
Tungsten	5.6×10^{-8}
Graphite	$(3-60) \times 10^{-5}$
Silicon	$0.1 - 60$
Hard Rubber	$(1-100) \times 10^{13}$

Definition 22.1.4. The power absorbed by a resistor can be calculated according to

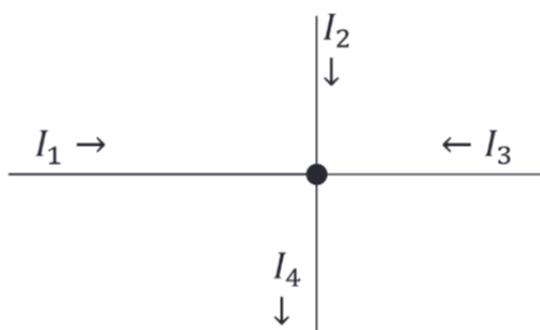
$$P(t) = v(t)i(t) = i^2(t)R = \frac{v^2(t)}{R}$$

For DC current this reduces to $P = IV = I^2R = \frac{V^2}{R}$.

22.1.2 Circuit Laws

Theorem 22.1.1 (Kirchhoff's Current Law). The algebraic sum of all the currents entering and leaving a node is equal to zero.

$$\sum_{i=1}^n I_i = 0$$



This is really conservation of charge.

Theorem 22.1.2 (Kirchhoff's Voltage Law). *The algebraic sum of all the voltage drops around any closed path is zero at every distance and time:*

$$\sum_{i=1}^n V_i = 0$$

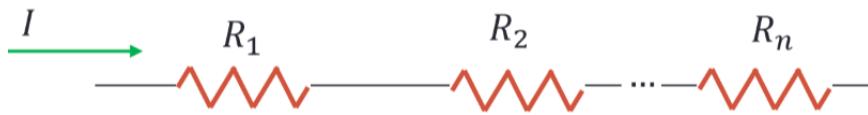
This is a form of energy conservation.

For a circuit and any pair of nodes j and k , the voltage drop v_{jk} from node j to node k is

$$v_{jk} = v_j - v_k$$

Theorem 22.1.3 (Resistors in Series). *If we have a line of n resistors in series (i.e. the same current passes through all of them), then the equivalent resistance of the collection is their sum:*

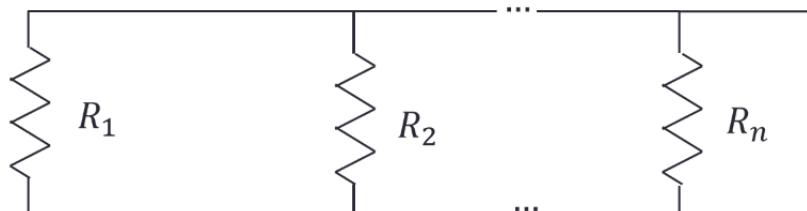
$$R_{eq}^{series} = \sum_{i=1}^n R_i$$



Now, consider a circuit that has n resistors in series connected to a battery, V_{in} , with current I . Then the equivalent resistance across the series of resistors is $R_{eq} = V_{in}/I$, so the voltage drop across any individual resistor V_i is

$$V_i = R_i I = R_i \frac{V_{in}}{\sum_{j=1}^n R_j}$$

Definition 22.1.5. *n resistors are said to be in parallel if the same voltage drop happens across them:*



In this case different currents go through each resistor, say current I_i goes through resistor R_i , and we have

$$I_T = \frac{V_{in}}{R_{eq}} = \sum_{i=1}^n I_i = \sum_{i=1}^n \frac{\Delta V_i}{R_i}$$

where by Kirchhoff's voltage law $V = \Delta V_i$ for each I , so

$$\frac{1}{R_{eq}} = \sum_{i=1}^n \frac{1}{R_i}$$

22.1.3 Voltmeters and Ammeters

Definition 22.1.6. A voltmeter measures voltage (potential difference) between two points along the circuit. We connect them in parallel to the device. To connect them we create a node, which means some current will always go through the resistor. The goal is to send as little current as possible through the voltmeter, therefore voltmeters have very high internal resistances.

Definition 22.1.7. Ammeters measure current through the circuit. We connect them in series with other elements. In order not to affect the circuit, we want them to have the smallest internal resistance possible.

22.1.4 Thévenin Theorems

Theorem 22.1.4. Given an arbitrary two-terminal linear network N consisting of resistances and independent sources, then, for almost all such N , there exists an equivalent two-terminal network consisting of a resistance R_{th} in series with an independent voltage source $V_{oc}(t)$.

The voltage $V_{oc}(t)$ (open-circuit voltage) is what appears across the two terminals of N when no other network is attached. Resistance R_{th} (Thévenin Equivalent Resistance) is the equivalent resistance of N when all independent sources are deactivated.

The procedure for calculations is as follows:

1. Remove the load resistor R_L or component concerned. (Create an opening between two terminals)
2. Remove the active sources: Remove voltage sources by shortening them, and remove current sources by opening them
3. Find R_{th} using rules for regular equivalent resistances.
4. Find V_{oc} by the usual circuit analysis methods.
5. Find the current flowing through the load resistor R_L .

22.2.0 RLC Circuits

22.2.1 Phasors

First recall we can consider an oscillation $s(t) = A \cos(\omega t + \phi_0)$ as the real part of a complex function

$$\tilde{s}(t) = Ae^{i(\omega t + \phi_0)}$$

Similarly, $y(t) = A \sin(\omega t + \phi_0)$ can be seen as the imaginary part of the same function. $\tilde{s}(t)$ is our **phasor**. Note that waves add as phasors, which add vectorially (superposition principle). In general, any oscillation can be broken into a sum of simple harmonic oscillations (i.e. Fourier series).

22.2.2 Capacitors

Now, the capacitor usually consists of a pair of parallel plates (possibly with a dielectric between them) storing opposite charges on either side, and hence containing a potential difference between the plates.

Definition 22.2.1. **Capacitance** is the ratio of the magnitude of the charge on one of the plates of the capacitor to the absolute value of the potential difference between them is defined as

$$C = \frac{q}{|\Delta V|}, \quad 1 \text{ F} = \frac{1 \text{ C}}{1 \text{ V}}$$

For a parallel-plate capacitor of area A and plate separation d we can compute the capacitance by applying Gauss's Law $\oint \vec{E} \cdot d\vec{A} = \frac{q_{in}}{\epsilon_0}$, by

$$\oint \vec{E} \cdot d\vec{A} = EA = \frac{q}{\epsilon_0} = \frac{\sigma A}{\epsilon_0}$$

$$\text{so } E = \frac{\sigma}{\epsilon_0}, \text{ so } \Delta V = Ed = \frac{\sigma d}{\epsilon_0} = \frac{qd}{A\epsilon_0}, \text{ and } C = \frac{A\epsilon_0}{d}.$$

In order to determine the electric potential energy stored in a capacitor, we look at energy required to transfer charge from one conductor to another. Consider moving an infinitesimal amount of charge dq' , so the potential difference is essentially constant during the transfer of one increment $dU^E = -dW = Vdq' = \frac{q'}{C}dq'$, so $U^E = \frac{1}{2} \frac{q^2}{C} = \frac{1}{2}C\Delta V^2 = \frac{1}{2}q\Delta V$.

22.2.3 Inductor

A consequence of Faraday's Law, $\oint \vec{E} \cdot d\vec{l} = -\frac{d\Phi_B}{dt}$, where Φ_B is the magnetic flux, is that when a current through a conducting loop changes it induces an emf (potential difference, ϵ_{ind}) in the loop itself. **Inductance** L of the loop (or solenoid) is defined as a proportionality constant

between the emf and the rate of change of current:

$$\varepsilon = -L \frac{dI}{dt}, \quad 1 \text{ H} = \frac{1 \text{ Vs}}{1 \text{ A}}$$

A device that has an appreciable inductance is called an **inductor**. It describes the change of magnetic flux associated with a change of current in a particular loop/solenoid.

Work must be done on an inductor to establish the current through it because the change in current causes an induced emf that opposes this change. Consider dW , work done *on* the inductor when an amount of charge dq is being moved through it, so $dW = -\varepsilon dq$, $\frac{dW}{dt} = -\varepsilon I = LI \frac{dI}{dt}$, and so $W = \int dW = \frac{1}{2}LI^2$ and $U^B = \frac{1}{2}LI^2$.

22.2.4 AC Current

For a conductor, at any instance $C = \frac{q}{v_c(t)}$ so $q = Cv_c(t) = CV_c \sin(\omega t)$. Current is

$$i(t) = CV_c \omega \cos(\omega t) = CV_c \omega \sin\left(\omega t + \frac{\phi}{2}\right) = I_c \sin\left(\omega t + \frac{\phi}{2}\right)$$

As $I_C = CV_C \omega$, $V_C = \frac{I_C}{C\omega}$, so $Z_C = \frac{1}{\omega C}$.

First, $v_L(t) = L \frac{di}{dt} = V_L \sin(\omega t)$, so

$$i(t) = -\frac{V_L}{\omega L} \cos(\omega t) = \frac{V_L}{\omega L} \sin\left(\omega t - \frac{\pi}{2}\right)$$

Then

$$I_L = \frac{V_L}{\omega L} \implies Z_L = \omega L$$

22.2.5 Series RLC Circuit

First, for a circuit consisting of an inductor, resistor, capacitor, and an AC voltage source, we have the relation $\varepsilon = v_R(t) + v_L(t) + v_C(t)$. Current is the same for all three elements, so

$$V_0 \sin(\omega t) = \frac{1}{C} \int idt + L \frac{di}{dt} + iR$$

and differentiating

$$V_0 \omega \cos(\omega t) = \frac{i}{C} + L \frac{d^2i}{dt^2} + R \frac{di}{dt}$$

The solution to these DEs is

$$i(t) = \frac{V_0 \sin(\omega t - \phi)}{\sqrt{R^2 + (\omega L - \frac{1}{\omega C})^2}}$$

where

$$\phi = \tan^{-1} \left[\frac{\omega L - \frac{1}{\omega C}}{R} \right]$$

Chapter 23

AC Sources, Waveforms

Recall that “AC” stands for *alternating current* - current whose value oscillates sinusoidally about a mean value of zero through time. Recall also from Fourier transform theory that any well-behaved signal of any shape can be synthesized as a linear combination of pure AC signals of appropriately chosen frequencies, amplitudes, and phases.

First, in a circuit diagram we notate an AC voltage source by:

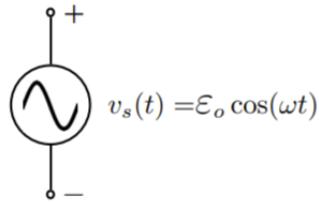


Figure 23.1: Circuit symbol for an ideal AC voltage source. \pm signs show the polarity of the output voltage at $t = 0$.

The defining property of an ideal AC voltage source is that it produces a sinusoidally varying voltage signal of constant amplitude and frequency across its terminals, no matter what device is connected across them

$$v_s(t) = v_+ - v_- = \varepsilon_0 \cos(\omega t)$$

where ε_0 and ω are positive real constants. ε_0 is the constant *voltage amplitude* of the source, measured in volts (V) SI units. ω is the *angular frequency* of the source, measured in radians per second. The waveform executes one complete cycle in a time interval equal to the period T :

$$T = \frac{1}{f} = \frac{2\pi}{\omega}$$

This is illustrated in the Figure:

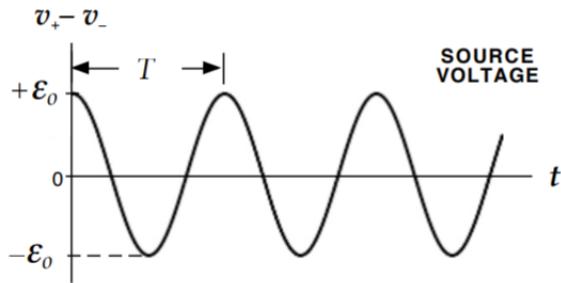


Figure 23.2: Output voltage signal of the ideal AC voltage source of Figure ??.

Often we will ground the lower terminal of an ideal voltage source, so that all potentials in the circuit are measured relative to that ground.

Definition 23.0.1. The root-mean-square voltage, or V_{rms} , of an AC voltage source is the square root of the time-averaged square of the voltage signal. For the AC voltage source this is just

$$V_{rms} = \frac{1}{\sqrt{2}}\varepsilon_0$$

Sometimes the amplitude will be given in terms of the rms voltage. A third way of specifying the voltage amplitude is the peak-to-peak voltage swing, or maximum voltage minus minimum voltage for one cycle of the voltage waveform. For an AC waveform this is simply twice the amplitude.

23.1.0 Loaded Ideal AC Voltage Generator

Consider the AC circuit in which an ideal AC voltage source of the type described previously is connected across a circuit element, referred to as the load. The load is a “black box” with two terminals. If the load contains only some combination of inductance, resistance, and capacitance, then it is said to be a passive linear load.

Proposition 23.1.1. When considering an AC circuit with a passive linear load as the only component, the laws of physics dictate that the current $i(t)$ in the circuit must be a constant-amplitude sinusoidal function of time with the same angular frequency as the ideal voltage source:

$$i(t) = I_0 \cos(\omega t + \phi)$$

I_0 is the current amplitude, measured in amperes (A), and ϕ is the phase angle.

When the load is pure resistance R , the law of the resistor dictates that $v_s = iR$ (Ohm’s Law). Thus, in the resistive case, $\phi = 0$, the current has amplitude $I_0 = \frac{\varepsilon_0}{R}$, and the current is exactly in phase with the source voltage $v_s(t)$. The presence of capacitors and/or inductors is what gives us a phase shift.

23.2.0 Three Fundamental Circuits

An important problem in AC circuit theory is finding the current in the circuit when the load network is specified. We shall first investigate the simple case where the load consists of a single fundamental circuit element: a resistor, a capacitor, or an inductor. In the case of a resistor, we have the following law:

Theorem 23.2.1 (Law of the Resistor). *The voltage across a resistor is given by*

$$v_R(t) = i(t)R$$

Putting $v_R = v_s(T) = \varepsilon_0 \cos(\omega t)$, we obtain

$$i(t) = \frac{\varepsilon_0}{R} \cos(\omega t)$$

Thus, we have current amplitude $I_0 = \frac{\varepsilon_0}{R}$ and current phase $\phi = 0$.

For a capacitor we have the following law:

Theorem 23.2.2 (Law of the Capacitor). *The voltage across a capacitor with capacitance C and storing charge $q(t)$ is*

$$v_C(t) = \frac{q(t)}{C}$$

or

$$i(t) = \frac{dq}{dt} = C \frac{dv_C}{dt}$$

With $v_C = v_s(t) = \varepsilon_0 \cos(\omega t)$, $i(t)$ is

$$i(t) = -\omega C \varepsilon_0 \sin(\omega t) = \omega C \varepsilon_0 \cos(\omega t + \pi/2)$$

The current amplitude is $I_0 = \omega C \varepsilon_0$, and the current phase is $\phi = \pi/2$.

Finally, the law for inductors is as follows:

Theorem 23.2.3 (Law of the Inductor). *The voltage across an inductor with inductance L is given by*

$$v_L(t) = L \frac{di(t)}{dt}$$

or

$$i(t) = \frac{1}{L} \int_0^t v_L(t') dt'$$

Putting $v_L = v_s(t) = \varepsilon_0 \cos(\omega t)$ and solving for $i(t)$ gives

$$i(t) = \frac{\varepsilon_0}{\omega L} \sin(\omega t) = \frac{\varepsilon_0}{\omega L} \cos(\omega t - \pi/2)$$

The current amplitude is $I_0 = \frac{\varepsilon_0}{\omega L}$ and the current phase is $\phi = -\pi/2$.

We can summarise these results as follows:

- When an AC voltage signal of amplitude V_0 and angular frequency ω appears across any linear circuit element, the current in the element is guaranteed to be an AC current signal of the same angular frequency ω as the voltage signal.
- When the device is a **resistor** of resistance R , the current is exactly in phase with the voltage across the resistor ($\phi = 0$), and the current amplitude is given by $I_0 = V_0/R$.
- When the device is a **capacitor** of capacitance C , the current is $\pi/2$ radians ahead of the voltage across the capacitor ($\phi = \pi/2$); the current amplitude is frequency dependent and given by $I_0 = \omega C V_0$
- When the device is an **inductor** of inductance L , the current is $\pi/2$ radians behind the voltage across the inductor ($\phi = -\pi/2$); the current amplitude is frequency dependent and given by $I_0 = V_0/(\omega L)$

Note that in the case of a conductor, the source voltage directly determines the charge on the upper plate of the capacitor, not the current, as was the case with the resistive load. The current is then the resulting time derivative of the capacitor charge q . Indeed, the current is the scaled derivative of the source voltage.

Chapter 24

Phasor Analysis

A **phasor** is simply a complex number, a constant independent of time, that contains information on both the amplitude and phase of the AC signal.

24.1.0 Phasors and Complex Sinusoids

Recall that we can use Euler's relation to express a complex exponential as

$$e^{j\theta} = \cos \theta + j \sin \theta$$

where $j = \sqrt{-1}$, and that the modulus of $e^{j\theta}$ is 1. If we set $j\theta = j\omega t$ we will have made what is called a time-dependent **complex sinusoid of unit amplitude**, $e^{j\omega t}$.

Recall that for a single loop circuit, one with a linear load, we have the current

$$i(t) = I_0 \cos(\omega t + \phi)$$

Definition 24.1.1. We define the time-independent **complex amplitude** or **phasor** associated with $i(t)$ as

$$\hat{I} = I_0 e^{j\phi}$$

Combining the phasor \hat{I} with the complex sinusoid $e^{j\omega t}$ we can express the current as

$$i(t) = \operatorname{Re}(\hat{I} e^{j\omega t}) = I_0 \cos(\omega t + \phi)$$

Using our phasor \hat{I} , we can recover the amplitude and phase of the physical current as follows:

$$I_0 = |\hat{I}| = \sqrt{(\operatorname{Re}\hat{I})^2 + (\operatorname{Im}\hat{I})^2}; \quad \phi = \tan^{-1} \left[\frac{\operatorname{Im}\hat{I}}{\operatorname{Re}\hat{I}} \right]$$

24.2.0 Kirchhoff's Rules

Theorem 24.2.1 (Kirchhoff's Point Rule). *For any junction in a circuit with n currents going into it, $i_1(t), \dots, i_n(t)$, where $i_j(t) < 0$ if it is conventional current going out of the junction, and $i_j(t) > 0$ if it is conventional current going into the junction, then*

$$\sum_{j=1}^n i_j(t) = 0$$

This is a statement of conservation of charge.

If this junction is part of a circuit network driven by a single AC voltage source of angular frequency ω , each of the currents will itself be a sinusoidal function of time with the same angular frequency ω as the source. Using complex phasor amplitudes, $\hat{I}_k = I_k e^{j\phi_k}$, we have that

$$\operatorname{Re} \left[\sum_{k=1}^n \hat{I}_k e^{j\omega t} \right] = 0$$

First, note that $\sum_{k=1}^n \hat{I}_k = \sum_{k=1}^n I_k e^{j\phi_k} = I_S e^{j\phi_S}$ for some I_S and ϕ_S since all complex numbers can be expressed in polar form. Then $\operatorname{Re}[I_S e^{j\phi_S} e^{j\omega t}] = 0$. Since this holds for all t , we must have that $I_S = 0$ (setting $t = \frac{-\phi_S}{\omega}$), and it follows that $I_S e^{j\phi_S} e^{j\omega t} = 0$. Thus we have the complex equation

$$\sum_{k=1}^n \hat{I}_k e^{j\omega t} = \sum_{k=1}^n I_k e^{j\phi_k} e^{j\omega t} = 0$$

Since $e^{j\omega t} \neq 0$, this is equivalent to the following:

Theorem 24.2.2 (Kirchhoff's point rule (phasors)). *Kirchhoff's point rule is equivalent to*

$$\sum_{k=1}^n \hat{I}_k = \sum_{k=1}^n I_k e^{j\phi_k} = 0$$

where \hat{I}_k is the phasor of the k th current going into the junction.

Next we have Kirchhoff's loop rule:

Theorem 24.2.3 (Kirchhoff's Loop Rule). *If one encounters n voltage drops/gains, $v_k(t)$, when going around a closed loop with zero resistance wires, Kirchhoff's loop rule state*

$$\sum_{k=1}^n v_k(t) = 0$$

This is a statement of the law of conservation of energy.

If the voltage drops are AC signals of phasor amplitudes $\hat{V}_k = V_k e^{j\phi_k}$, the loop rule can be transformed as with the junction rule:

Theorem 24.2.4 (Kirchhoff's Loop Rule (phasors)). Kirchhoff's loop rule is equivalent to

$$\sum_{k=1}^n \hat{V}_k = \sum_{k=1}^n V_k e^{j\phi_k} = 0$$

where \hat{V}_k is the phasor of the k th voltage drop/gain, if all voltage drops/gains are AC signals with the same angular frequency ω .

24.3.0 Phasor Summary

We summarize the properties of phasors as follows:

- A phasor is a complex constant (no time dependence) associated with a voltage or current waveform in an AC circuit.
- Any complex number $z = x + jy$ can be written in polar form $z = |z|e^{j\phi}$ with ϕ its phase constant or argument, and $|z|$ its modulus or magnitude.
- For an AC current signal $i(t) = I_{max} \cos(\omega t + \phi)$, the associated phasor is $\hat{I} = I_{max}e^{j\phi}$. If we have the phasor \hat{I} we can recover the amplitude and phase of $i(t)$ as $I_{max} = |\hat{I}|$ and $\phi = \arg \hat{I}$. Similarly, for any AC voltage signal $v(t) = V_{max} \cos(\omega t + \theta)$, the associated phasor is $\hat{V} = V_{max}e^{j\theta}$. If we have the phase \hat{V} we can recover the amplitude and phase of $v(t)$ as $V_{max} = |\hat{V}|$ and $\theta = \arg \hat{V}$.
- When phasors add, their real and imaginary parts add to produce the phasor sum.

Recall that complex numbers behave exactly like two-dimensional real vectors under addition. We can use this fact to draw vector addition diagrams, called phasor addition diagrams in our context.

24.4.0 Complex Impedance

Consider a circuit consisting of an ideal AC voltage source of frequency f connected across a load containing nothing but capacitors, resistors, and inductors, connected to each other in some arbitrary fashion. Such a load is called a passive linear load, since it contains no sources of emf. Recall that in this case $v(t) = \varepsilon_0 \cos \omega t$ and $i(t) = I_{max} \cos(\omega t + \phi)$. The phasors, \hat{V} and \hat{I} , representing the load are $\hat{V} = \varepsilon_0$ and $\hat{I} = I_{max}e^{j\phi}$.

Definition 24.4.1. The complex impedance Z of a load is the ratio of its voltage phasor to its current phasor:

$$Z = \frac{\hat{V}}{\hat{I}}$$

The unit of Z is the ohm ($1 \Omega = 1 \text{ V/A}$)

The impedance of a passive load is generally complex and frequency-dependent. We can always write the impedance in rectangular form

$$Z = R + jX$$

The real part and imaginary part of the impedance Z are both purely real quantities with units of ohms. They are referred to as the ***resistance*** (R), and ***reactance*** (X) of the load.

Given the source voltage phasor and the impedance Z of the load, we can find the complex current phasor using the form of Z . The amplitude and phase relative to the voltage source of the current are then

$$I_{max} = |\hat{I}| = \frac{\varepsilon_0}{|Z|}, \text{ and } \phi = \arg \hat{I} = -\arg Z$$

Observe that through this process, we have simplified the fundamental laws by Kirchhoff and Ohm for AC circuits to ones purely involving complex numbers, giving the same form as the DC case.

Chapter 25

Impedances in Series

We wish to find impedance expressions for circuits with pure resistance, pure capacitance, and pure inductance.

25.1.0 Impedances of Simple Linear Devices

First we review our time-dependent relations for current of simple circuits with only resistance, capacitance, or inductance, as well as their associated phasor relations in the following table:

Table 25.1: Voltage-current relations in the time and phasor domains (assumes a voltage waveform $v(t) = V_{max} \cos \omega t$ across the load)

Load type	time-dependent relation	phasor relation
Resistive load (R)	$i(t) = \frac{V_{max}}{R} \cos \omega t$	$\hat{I} = \frac{\hat{V}}{R}$
Capacitative load (C)	$i(t) = C\omega V_{max} \cos(\omega t + \pi/2)$	$\hat{I} = C\omega \hat{V} e^{j\pi/2} = j\omega C \hat{V}$
Inductive load (L)	$i(t) = \frac{V_{max}}{\omega L} \cos(\omega t - \pi/2)$	$\hat{I} = \frac{\hat{V}}{\omega L} e^{-j\pi/2} = \frac{\hat{V}}{j\omega L}$

Note that our definition of complex impedance gives $\hat{I} = \frac{\hat{V}}{Z}$. We collect these results in the following table:

Table 25.2: Impedances of basic circuit elements (ω = angular frequency of source)

circuit element	impedance, Z
Resistor (R)	R
Capacitor (C)	$\frac{1}{j\omega C}$
Inductor (L)	$j\omega L$

25.2.0 The Voltage Divider

Note that many results from DC circuit theory can be applied to AC circuit theory in the phasor domain. The addition law for a series connection of impedances, for instance, is impedances in series add to give the equivalent impedance:

$$Z_{eq} = \sum_{i=1}^n Z_i$$

Note that a string of impedances is considered to be connected in series only when the impedances carry exactly the same current. The common current in the series combination is given by

$$\hat{I} = \frac{\hat{V}}{Z_{eq}} = \frac{\hat{V}}{\sum_{i=1}^n Z_i}$$

where \hat{V} is the voltage phasor for the voltage source. The source voltage is divided across the series of impedances so that the voltage across the kth element in the series is proportional to its own impedance:

$$\hat{V}_k = \hat{I}Z_k = \frac{\hat{V}}{Z_{eq}}Z_k = \frac{Z_k}{\sum_{i=1}^n Z_i}\hat{V}$$

Note 25.2.1. Note that in the last equation, \hat{V}_k , Z_k and the constant of proportionality \hat{I} are all complex quantities in general. In other words we are adding phasor voltage drops \hat{V}_k to get the source voltage rise \hat{V} . Kirchhoff's loop rule dictates that the phasor sum of the voltage drops is equal to the source voltage phasor

$$\sum_{i=1}^N \hat{V}_i = \hat{V}$$

and therefore the magnitudes of the phasor sum and the source voltage must be the same

$$\left| \sum_{i=1}^N \hat{V}_i \right| = |\hat{V}|$$

but the sum of the individual amplitudes of the sources voltages generally exceeds the amplitude of the phasor sum

$$\sum_{i=1}^N |\hat{V}_i| \geq |\hat{V}|$$

Equality holds only if all the phasors in the summation are collinear and pointing in the same direction. This only occurs if all the elements in the combination have impedances with the same complex phase.

25.3.0 The RC Series Circuit

We now investigate the RC series circuit, depicted in Figure ??:

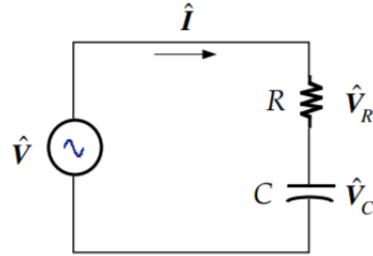


Figure 25.1: RC series circuit.

We assume the source voltage phasor is purely real and equal to \hat{V}_S ; equivalently, we will measure the phases of all oscillating quantities relative to the phase of the source voltage, which has phase zero. We first obtain the equivalent complex impedance of the RC series combination in our circuit:

$$Z_{eq} = Z_R + Z_C = R + \frac{1}{j\omega C} = \frac{1 + j\omega RC}{j\omega C}$$

Then the current phasor in the circuit is found using the definition of impedance:

$$\hat{I} = \frac{\hat{V}}{Z_{eq}} = \frac{j\omega C}{1 + j\omega RC} \hat{V}_S$$

The amplitude and phase of the current phasor are then given by

$$|\hat{I}| = \frac{\omega C}{\sqrt{1 + \omega^2 R^2 C^2}} V_S$$

and

$$\phi = \arg \hat{I} = \arg(j\omega C) - \arg(1 + j\omega RC) + \arg V_S = \frac{\pi}{2} - \tan^{-1}(\omega RC)$$

Now we can find the phasor voltage dropped across the resistor and capacitor by using the first form of the voltage divider equations:

$$\hat{V}_R = \hat{I}R = \frac{j\omega RC}{1 + j\omega RC} \hat{V}_S$$

and so

$$|\hat{V}_R| = \frac{\omega RC}{\sqrt{1 + \omega^2 R^2 C^2}} V_S$$

and

$$\arg \hat{V}_R = \arg \hat{I} + \arg R = \arg \hat{I} = \frac{\pi}{2} - \tan^{-1}(\omega RC)$$

Similarly, for the capacitor,

$$\hat{V}_C = \hat{I}Z_C = \frac{\hat{I}}{j\omega C} = \frac{1}{1 + j\omega RC} \hat{V}_S$$

with amplitude

$$|\hat{V}_C| = \frac{1}{\sqrt{1 + \omega^2 R^2 C^2}} V_S$$

and phase

$$\arg \hat{V}_C = \arg \hat{I} + \arg Z_C = \arg \hat{I} - \pi/2 = -\tan^{-1}(\omega RC)$$

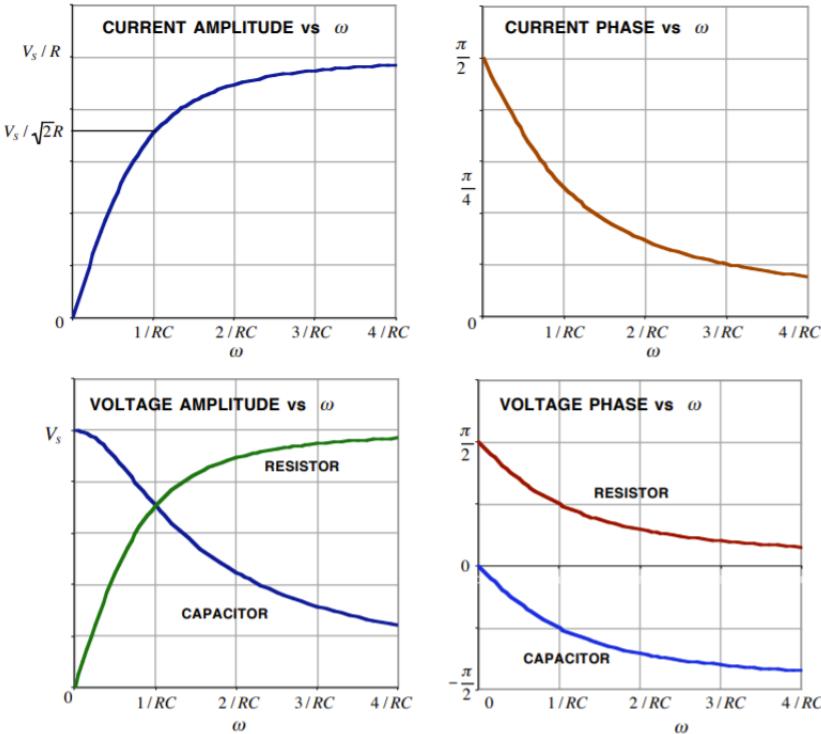


Figure 25.2: Graphs of current and voltage amplitudes and phases as a function of angular frequency.

At low frequencies, $\omega \ll 1/RC$, the impedance $1/j\omega C$ of the capacitor is high, keeping the current amplitude small. The amplitude of the capacitor voltage is close to the source amplitude V_S at these low frequencies. At high frequencies, $\omega \gg 1/RC$, the capacitor impedance is small and the amplitude of the current is limited by the resistor to an asymptotic value of V_S/R ; the voltage drop across the resistor is approximately V_S , with the capacitor voltage drop being very small. At all frequencies the voltage signal dropped across the resistor always leads the voltage signal dropped across the capacitor by $\pi/2$ of phase. Since the phase difference is $\pi/2$, and we have the phasor relation $\hat{V}_S = \hat{V}_R + \hat{V}_C$, we have that \hat{V}_R and \hat{V}_C are orthogonal so pythagoras' theorem applies and $V_S^2 = V_C^2 + V_R^2$.

Remark 25.3.1 (Analog filters). The RC circuit is useful as an analog filter to remove unwanted high or low frequency components from a time-continuous signal that contains a spectrum of frequencies. The filter can be said to have a characteristic angular frequency $\omega_C = 1/RC$ that can be seen as separating low frequencies from high frequencies. If we measure an output voltage signal across the resistor, we find that we strongly reject frequencies lying well below ω_C and retain those that lie well above ω_C . This is called a high pass filter. On the other hand, if we measure the output signal across the capacitor, we will have a low pass filter which also has its uses. Serious filter design usually requires cascading a number of RC or RLC circuits in a carefully designed series to get reasonable filter performance. Such filters typically rely on inserting operational amplifiers in the cascade for successful implementation.

Remark 25.3.2 (Transient versus Steady State). The ideas of AC circuits that we have developed assume that the AC source that drives the circuit has been connected for a time sufficiently long enough for exponentially decaying signals, called transients, to have died away, leaving only the simple steady state sinusoidal AC signals which have constant amplitude and phase,

and which will persist unchanged until the source is disconnected at some future time, after which new transients emerge which bring down the steady state outputs smoothly and exponentially to zero. When we are thinking in terms of analog filters, we are thinking of the steady state response of the RC circuit.

Chapter 26

The Series RLC Circuit

26.1.0 Differential Equation for the RLC circuit

The RLC circuit is an AC circuit consisting of a capacitor, resistor, and inductor all connected in series:

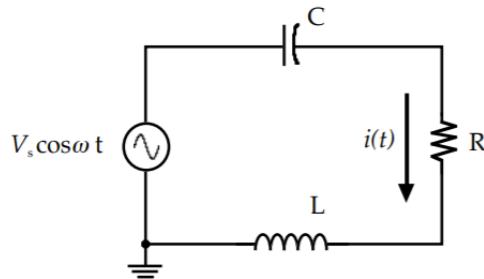


Figure 26.1: The series RLC circuit, driven by an AC voltage source.

Applying Kirchhoff's loop rule (in the time domain) to the series RLC circuit in Figure ?? gives the following equation expressed in terms of the time-varying charge $q = \int idt$ stored in the capacitor:

$$L \frac{d^2q}{dt^2} + R \frac{dq}{dt} + \frac{q}{C} = V_s \cos \omega t$$

Note that this differential equation takes the form of a damped harmonic oscillator with a harmonic driving term:

$$\ddot{x} + 2\gamma \dot{x} + \omega_0^2 x = A \cos \omega t$$

The damping constant for us is $\frac{R}{2L}$, and the natural frequency of the undamped oscillator is $\omega_0 = \frac{1}{\sqrt{LC}}$. The RLC circuit also exhibits resonance (a large current amplitude when the driving frequency ω is equal to the natural frequency $\omega_0 = 1/\sqrt{LC}$).

26.2.0 Resonance in the Series RLC Circuit

We assume that the voltage source of Figure ?? has been connected for a long time so that any transient currents have died away, and the AC current in the circuit has reached its steady state. In terms of voltage and current phasors we can write our DE as

$$\hat{I}(Z_C + Z_R + Z_L) = \hat{V}_s$$

where the impedance terms are $Z_C = 1/j\omega C$, $Z_R = R$, $Z_L = j\omega L$. Assigning a phase of zero to the source voltages, thus making \hat{V}_s purely real, results in the following expression for the complex current amplitude in the circuit:

$$\hat{I} = \frac{V_s}{\frac{1}{j\omega C} + R + j\omega L} = \frac{V_s}{R + j[\omega L - \frac{1}{\omega C}]}$$

The real amplitude I and phase ϕ of the AC current are then

$$I = \frac{V_s}{\sqrt{R^2 + [\omega L - \frac{1}{\omega C}]^2}}; \quad \phi = -\tan^{-1} \left[\frac{\omega L}{R} - \frac{1}{\omega RC} \right]$$

Note the current amplitude has a maximum, with respect to ω , when $\omega = \omega_0 = \frac{1}{\sqrt{LC}}$. Thus, the natural angular frequency ω_0 of the undamped oscillator is also the resonance frequency of the RLC circuit. At resonance, $\omega = \omega_0$, the current amplitude achieves its maximum value of V_s/R . Also the value of ϕ is zero at resonance; this means that the current is in phase with the source voltage when the current amplitude is maximum.

Since the current amplitude at resonance is proportional to $1/R$, changing R while holding L and C fixed will change the height of the peak of the resonance curve without changing the resonance angular frequency, $\omega_0 = 1/\sqrt{LC}$.

One useful time scale in the RLC circuit is $1/\omega_0 = \sqrt{LC}$. Another comes from defining a dimensionless parameter Q called the quality factor of the circuit:

$$Q = \frac{\sqrt{L/C}}{R}$$

Circuits of given L and C with high Q will have small values of R . Low- Q circuits will have high values of R . We now define the normalized angular frequency $\bar{\omega}$:

$$\bar{\omega} = \frac{\omega}{\omega_0} = \omega \sqrt{LC}$$

Using these definitions in the current amplitude and phase equations we can write

$$I = \frac{V_s/R}{\sqrt{1 + Q^2 [\bar{\omega} - \frac{1}{\bar{\omega}}]^2}}; \quad \phi = -\tan^{-1} \left[Q \left(\bar{\omega} - \frac{1}{\bar{\omega}} \right) \right]$$

At resonance, the reactance of the RLC circuit disappears, and the circuit acts as a pure resistance (the impedance becomes purely real). The source voltage is dropped entirely across the resistor, and the current amplitude is V_s/R irrespective of the values of L and C . It can be seen that increasing Q for given values of L and C , by decreasing the resistance R in the circuit, will sharpen the resonance curve.

26.3.0 Voltage Characteristics

We now wish to look at the AC voltage signals that appear across each of the elements in the RLC circuit, as represented by their three phasors:

$$\begin{aligned}\hat{V}_R &= \hat{I}Z_R = \hat{I}R \\ \hat{V}_C &= \hat{I}Z_C = \hat{I} \frac{1}{j\omega C} \\ \hat{V}_L &= \hat{I}Z_L = \hat{I}j\omega L\end{aligned}$$

The phasor $\hat{I} = I(\omega)e^{j\phi(\omega)}$ is common to all three expressions. It follows that

- The AC voltage signal dropped across the **resistor** has amplitude $V_R(\omega) = I(\omega)R$ and phase $\phi_R = \phi(\omega)$ (because R is a positive real quantity)
- The AC voltage signal across the **capacitor** has amplitude $V_C(\omega) = I(\omega)/\omega C$ and phase $\phi_C = \phi(\omega) - \pi/2$ (because ωC is a positive real quantity and $1/j$ has phase $-\pi/2$)
- The AC voltage signal across the **inductor** has amplitude $V_L = \omega L I(\omega)$ and phase $\phi_L = \phi(\omega) + \pi/2$ (because ωL is a positive real quantity and j has phase $\pi/2$)

Using the results for $I(\omega)$ and $\phi(\omega)$ found previously we then have

$$\begin{aligned}V_R(\omega) &= \frac{V_s R}{\sqrt{R^2 + [\omega L - \frac{1}{\omega C}]^2}}; \quad \phi_R(\omega) = \phi(\omega) = -\tan^{-1} \left[\frac{\omega L}{R} - \frac{1}{\omega R C} \right] \\ V_C(\omega) &= \frac{V_s / \omega C}{\sqrt{R^2 + [\omega L - \frac{1}{\omega C}]^2}}; \quad \phi_C(\omega) = -\tan^{-1} \left[\frac{\omega L}{R} - \frac{1}{\omega R C} \right] - \frac{\pi}{2} \\ V_L(\omega) &= \frac{\omega L V_s}{\sqrt{R^2 + [\omega L - \frac{1}{\omega C}]^2}}; \quad \phi_L = -\tan^{-1} \left[\frac{\omega L}{R} - \frac{1}{\omega R C} \right] + \frac{\pi}{2}\end{aligned}$$

Chapter 27

Impedances in Parallel

Chapter 28

Models of Real AC Sources; Thévenin's Theorem

Chapter 29

Power Dissipation

Part III

Operational Amplifiers, Linear Systems, and the Fourier Transform

Chapter 30

Operational Amplifiers

Chapter 31

Linear Two-ports: Transfer Function, Impulse Response

Chapter 32

Time and Frequency Domains

Chapter 33

Data Sampling

Chapter 34

Analog Filter Design

Part IV

Semiconductors, Diodes, and Transistors

Chapter 35

Semiconductors

Chapter 36

Diodes

Chapter 37

Transistors

Part V

Laboratory Designs and Theory

37.A.0 Traveling Waves Theory

37.B.0 X-Ray Theory

37.C.0 Spectroscopy Theory

Appendices

.1.0 Uncertainty and Measurement

Measurement is a process that results in a value that can be attributed to a quantitative property of a phenomenon, object or substance. The object of the measurement (*measurand*) has the property that it has a numerical magnitude and a reference that gives its meaning (its *unit*). The process involves a comparison with an accepted standard, unless one does not yet exist. Instruments for measurement need to be calibrated so that they properly relate to the relevant standard.

Question .1.1. How do we know if our measurements are adequate?

In a scientific experiment, measurement uncertainty is critical for determining whether an experiment agrees with a theoretical prediction or not. The BIPM (Bureau International des Poids et Mesures) “Guide to the expression of uncertainty in measurement” (GUM) is a widely accepted standard.

Definition .1.1. *Uncertainty in a measurement is a parameter that characterizes the dispersion of values that could reasonably be attributed to the measured quantity.*

*The term “**standard uncertainty**” should be used to refer to uncertainties calculated according to the GUM guidelines.*

Uncertainty is a quantification of the doubt about the measurement result.

Remark .1.2. Note: Uncertainty is distinct from *error*, which is the (generally unknown) difference between the measured value and the ‘true value’ of the measured quantity. **Random errors** cause values to be randomly distributed above and below the true value. **Systematic errors** causes all measurements to be consistently over or underestimate.

Many factors may contribute to the uncertainty, including those arising from imperfect knowledge of systematic effects. Uncertainty creeps in many direction (see GUM 3.3.2). First we have limits to our *measurement devices*:

- Your measuring device has a limited *resolution*.
- Your measuring device has a *calibration uncertainty*. How close is a millimeter on your ruler to an exact millimeter? Are there other factors (“influence quantities”) such as temperature that might affect that uncertainty?

Even if your measuring device is arbitrarily precise, and well-calibrated, there are still *inherent errors* that cannot be avoided:

- Errors often arise due to *incomplete definition* of the quantity being measured, or due to *imperfect realization* of that definition. For example, your height varies considerably based on your posture. Similarly, exactly which points on the feet and ehad are being considered will also affect the measurement of your height. If those factors aren’t part of

the definition of your “height,” or if they aren’t perfectly controlled during the measurement, then they become a source of uncertainty.

- There may be systematic effects on the measurement that can only be corrected to a limited perfection. Systematic effects should always be corrected to the best of the experimenter’s ability during analysis. Nonetheless, there will always be a limit to how well you know the systematic effects, and on how well you can correct the results. This contributes to an uncertainty due to incomplete knowledge of the required value of the correction. Note: the measurement uncertainty reflects your uncertainty in applying that correction, not the magnitude of the correction. Explicitly apply the correction during the analysis, not when writing down the raw data, and always explain the correction in your report.

.1.1 Type A Uncertainty

Repeated measurements under apparently identical conditions often give variable results. There are limits on both repeatability (under exactly the same conditions in a short time) and reproducibility (under slightly different conditions or over a long time). The variance in repeated measurements gives a rigorous way to quantify many types of uncertainty.

Definition .1.2. Type A uncertainty is derived through repeated measurements and statistical methods.

Type B uncertainties include all methods of estimating uncertainty that are not included in Type A.

The following are common statistical methods used for estimating Type A uncertainties:

Definition .1.3. The arithmetic sample mean (average) of n independent samples X_1, \dots, X_n is

$$\bar{X} = \frac{\sum_{k=1}^n X_k}{n}$$

Definition .1.4. The best unbiased estimator for the standard deviation of a sample of n independent random variables X_1, \dots, X_n of sample mean \bar{X} is

$$s(X_k) = \sigma = \sqrt{\frac{\sum_{k=1}^n (X_k - \bar{X})^2}{n - 1}}$$

In excel, we can use STDEV.S()

Definition .1.5. The best estimate for the standard deviation, or standard uncertainty, of the mean \bar{X} of a sample of n independent random variables X_1, \dots, X_n is

$$u(\bar{X}) = \frac{s(X_k)}{\sqrt{n}}$$

From these formulations we observe that Type A uncertainties decrease as sample sizes tend to infinity, which is a statement of the central limit theorem.

If repeated measurements are impractical, but some sources of uncertainty simply cannot be estimated through repeated measurements, Type B uncertainties are used. For example, calibration uncertainty is a Type B uncertainty that cannot be revealed through repeated measurements without access to a reference standard. Similarly, imperfect corrections for systematic effects typically result in Type B uncertainties.

.1.2 Quantifying Standard Uncertainties

The Central Limit Theorem is the statement that in the limit as n goes to infinite for a sample of n independent identically distributed random variables X_1, \dots, X_n , the distribution of the sample means \bar{X} follows a Gaussian or normal distribution with mean equal to the true mean of the X_i , and uncertainty equal to σ/\sqrt{n} , for $\sigma = s(X_k)$.

In relation to the standard deviation for Type A uncertainties, we have the following rules of thumb:

Proposition .1.1. *One can expect the measurement error, the unknown difference between the measurement error and the true value, to be less than the standard deviation about 68% of the time. For two standard deviations we have a 95.4% certainty, and for three standard deviations we have about a 99.7% certainty.*

Note that uncertainties are not safety factors, nor confidence intervals. Often we want large than a 68% confidence, so we multiply the standard deviation by a coverage factor k to obtain a “expanded uncertainty” with a higher level of confidence. For example, the discovery of the Higgs Boson used a coverage factor of $k = 5$ (corresponding to a 99.99994% confidence)

.1.3 Type B Uncertainties

Most uncertainties are Type B, simply because Type A uncertainties are often impractical. For example, the manufacturer’s quoted uncertainty for a measurement device is a Type B uncertainty. Type B uncertainties should be honestly estimated to conform to the same definition of “uncertainty” as for Type A, with the same amount of confidence (for example 68%). Rigorously this requires estimating the probability distribution and then assessing its standard deviation.

When evaluating a Type B uncertainty you should think “what range am I about 68% sure that the measurement fallse in?” Evidently, this is not something you can typically estimate to high accuracy, but an honest experiment requires an honest effort to evaluate the Type B uncertainties. There is no rigorous recipe for how to estimate Type B uncertainties. There is no one right way at estimation, but here are some examples:

Example .1.1 (Digital Quantization). With a digital instrument, there is always an uncertainty

in the last digit. Assuming that every value in the range $[x - \Delta, x + \Delta]$ is equally likely, for Δ half the size of the unit of the last digit (i.e. half the size of the length between last digit values), we have a uniform distribution with uncertainty $u = (x + \Delta - (x - \Delta))^2/12 = \Delta/\sqrt{3}$. More approximately, 68% of the probability is contained in the range $x \pm 0.68\Delta$, giving $u \approx 0.68\Delta$.

Example .1.2 (Rulers and Analog Meters). This is a case where there is no “right way.” It will depend on how small the divisions on the ruler are, how good your eyesight is, and how closely you can align the object with the ruler.

Example .1.3 (Half-Division Example). It is common to see uncertainty reported as half the smallest division on an analog instrument or half the last digit of a digital instrument. While frequent, this is not actually part of the GUM standard. Uncertainty of half a division is about typical for many analog measurements, and slightly too large for most digital measurements. It is usually a reasonable starting point, but should be avoided as a hard fast rule.

Example .1.4 (Calibration Reports). Manufacturers may issue calibration reports for their instruments stating the uncertainty in its calibration. It may be stated as a standard uncertainty, but more often will be stated as a 95% confidence interval $x \pm \Delta$, or sometimes even a 99% confidence interval. In these cases we can use the standard rule of thumb to convert these to the standard uncertainty by taking either half or one third.

Example .1.5 (Systematic Effects). You should always correct measurements for known systematic effects. The uncertainty reflects the uncertainty in that correction, not its magnitude.

Example .1.6 (Event Counting). Counting the number of random events that occur in a given time is subject to statistical fluctuation. Some intervals will naturally have more events than others. This leads to an uncertainty in the measured event rate. The appropriate statistical distribution in this case is the Poisson distribution, which has a standard deviation equal to the square root of the mean. Therefore, if the observed count rate is n then its uncertainty is \sqrt{n} .

.1.4 Combining Uncertainties

Often we will have multiple sources of uncertainty, u_i , that contribute to an overall total uncertainty in a measurement. Assuming these uncertainties are uncorrelated, they can be combined into a combined standard uncertainty u_c , defined by

$$u_c = \sqrt{\sum_{i=1}^N u_i^2}$$

In practice, one uncertainty source often dominates over the others, in which case the combined uncertainty is essentially just the dominant uncertainty.

It is important not to “double-count” your uncertainties. The final uncertainty should be an honest assessment of your doubt about the measurement result, neither unrealistically large nor unrealistically small.

Sometimes the quantity of interest must be calculated from one or more measurements. In these cases we use a process of propagation of uncertainty.

Definition .1.6. If the final quantity r is a function of N independent uncorrelated variables, x_1, \dots, x_N , with uncertainties $u(x_1), \dots, u(x_N)$, then the propagated standard uncertainty is

$$u_c(r) = \sqrt{\sum_{i=1}^N \left(\frac{\partial r}{\partial x_i} \right)^2 u(x_i)^2}$$

.1.5 Reporting Uncertainties

Reporting a measurement with a precision that greatly exceeds its measurement uncertainty gives a false sense of accuracy, and should be avoided. Unless there are good reasons to do otherwise, uncertainties should be reported to one significant digit. Measurements and uncertainties should be reported in the same units, and with the same number of decimal places. Using scientific notation if appropriate is advised.

.1.6 Comparing a Measurement with an Expected Value

Experiments comparing a measured value x to an expected value μ are often called **hypothesis tests** in statistics, where we test $x = \mu$ by asking “what is the likelihood that x and μ differ by at least the amount observed?” That is, we ask what is the likelihood that $x = \mu$, and they differ by the amount observed. This is often known as the Null Hypothesis, and if the likelihood is very low, say less than 5% as standard, then the hypothesis is rejected and the values are said to disagree. In measured likelihood is called the p -value.

If the experiment is dominated by well-characterized Type A errors, then there is a mathematically rigorous way to test the hypothesis known as a “Student’s t-test”. However, experiments are often not dominated by Type A errors. Instead we often use some reasonable approximations. First calculate the normalized difference

$$t = \frac{x - \mu}{u_c(x)}$$

where $u_c(x)$ is the uncertainty in x . A large value of t means that x and μ are dissimilar, while a small value of t means that x is very similar to μ .

We then take the assumption that x follows a normal distribution. Although not rigorously justified, we have constructed $u_c(x)$ specifically to make this assumption as defensible as possible. Under the normal assumption, there is an $\approx 68\%$ chance of observing $|t| < 1$, an $\approx 95.4\%$ chance of observing $|t| < 2$, and an $\approx 99.7\%$ chance of observing $|t| < 3$. If $|t|$ is extremely small, there is a high chance something went wrong in the experiment, and you have overestimated your uncertainties.

If we want to compare two measurements with uncertainties, we can use the modified t value

$$t = \frac{x_1 - x_2}{\sqrt{u_c(x_1)^2 + u_c(x_2)^2}}$$

which is again an approximation so as to be reasonable for not just Type A uncertainties but Type B uncertainties as well.

.1.7 Line Fitting

Often we want to quantify a linear relationship between two quantities. **Line fitting** is the problem of finding a line that describes the data well. Intuitively, the problem is equivalent to minimizing the distances between the points and the line. The most commonly-used method is “least-squares” fitting, where the minimized quantity is the sum of the squares of the vertical distances between the data points and the fit line:

$$\chi^2 = \sum_{i=1}^n [y_i - (mx_i + b)]^2$$

for data points (x_i, y_i) and fit $y = mx + b$.

The closed form of the minimization for these parameters of the least squares fit is

$$m = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

and

$$b = \bar{y} - m\bar{x}$$

These solutions turn out to be the most probable line if the errors in y are normally-distributed and the errors in x are negligible. In Excel this is implemented using the LINEST function. The LINEST function takes in arguments

`LINEST(known - ys, known - xs, const = TRUE, stats = TRUE)`

If dealing with a general polynomial, we raise $known - xs$ to $\{...\}$ with the powers of x inside, or $known - xs^C COLUMN(\$a : \$n)$ for n the highest power. We can also apply *ln* to $known - xs$ for logarithmic fits. This command will return an array of values, the first row being coefficients, going from a_n to a_0 (left to right), and then uncertainties left to right in the second row, and the R^2 value in the third row.

When dealing with a nonlinear function, we often attempt to convert it to a linear form by changing the dependent/independent variables being graphed on either axis, and possibly taking logarithms if dealing with exponential relationships.

There is evidently an uncertainty associated with the values of the fitted coefficients m and b . We can define a Type A uncertainty in m based on the scatter of repeated measurements about a line. Intuitively, this makes sense: we should become less certain about m as the scatter of the points about the line increases. This Type A uncertainty is given by

$$u(m) = \frac{\sum_{i=1}^n [y_i - (mx_i + b)]^2}{(n - 2) \sum_{i=1}^n (x_i - \bar{x})^2} \quad u(b) = u(m) \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$$

This is also implemented as part of the LINEST function in Excel.

.1.8 Line Fitting with Measurement Uncertainty

The uncertainty estimate in the previous section has a few important restrictions:

-
- It is strictly Type-A, and cannot account for fundamentally Type-B uncertainties such as uncertainties in calibration or uncertainties in correcting for a systematic effect.
 - It weights all points equally, assuming they all have the same uncertainty
 - It assumes the x_i values are exact, attributing all uncertainty to y_i

These restrictions can affect both the uncertainties and the values of the m and b . We can lift the first two restrictions in the list by propagating uncertainty values through the linear lin-fitting equations, which allows for inclusion of Type-B uncertainties and unequal uncertainties for individual points.

The last restriction in the list comes from using the vertical distances instead of the uncertainty-weighted shortest distances in the definition of χ^2 . It becomes a valid assumption only if

$$u(y) \gg \frac{dy}{dx} u(x)$$

which is only sometimes true. If the converse is true, consider simply reversing x and y . If the uncertainties are comparable then you may combine the uncertainty in $u(x)$ with the uncertainty in $u(y)$

$$\sigma_i = \sqrt{u(y_i)^2 + \left[\frac{dy}{dx} u(x_i) \right]^2}$$

and then use σ in place of $u(y)$.

If the uncertainty in x , $u(x)$, is significant, then the combined uncertainty σ_i should be calculated using the last equation. Otherwise, let $\sigma_i = u(y_i)$, and then compute

$$\Delta = \sum_{i=1}^n \frac{1}{\sigma_i^2} \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right)^2$$

Next, compute m and b as

$$m = \frac{1}{\Delta} \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2} - \sum_{i=1}^n \frac{x_i}{\sigma_i^2} \sum_{i=1}^n \frac{y_i}{\sigma_i^2} \right)^2$$

$$b = \frac{1}{\Delta} \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \sum_{i=1}^n \frac{y_i}{\sigma_i^2} - \sum_{i=1}^n \frac{x_i}{\sigma_i^2} \sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2} \right)^2$$

and the uncertainties are

$$u(m) = \sqrt{\frac{1}{\Delta} \sum_{i=1}^n \frac{1}{\sigma_i^2}}$$

$$u(b) = \sqrt{\frac{1}{\Delta} \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}}$$

For lines fitted in this manner, data can be expected to lie within the range

$$y = mx + b \pm k \sqrt{[u(b)]^2 + [(x - \bar{x}_i)u(m)]^2}$$

.2.0 Graphing and Lab Reports

Tables should be numbered and captioned above. Figures are numbered and captioned below - explain the figure as a whole, what the data points represent, and what the line (if there is one) represents). That is explain all elements in the graph in its caption.

A standard lab report should have the following:

- **Cover Page:** Includes a title, names, UCIDs (School IDs), Section, Group number, TA's name, Date, and Location
- **Objective:** 2-3 sentences stating the purpose of the experiment/goals to be accomplished
 - “to determine...”
 - “to calculate...”
 - “to compare...”
 - “to investigate...”
- **Theory:** (*NOT* a textbook chapter) Should include the most important theory relevant to the experiment. Quote all equations you will use, and reference sources.
- **Experimental Methods (Methods and Materials):** Lists are welcome. Brief descriptions. Feel free to include sketches, and figures (with captions) of the experimental set up if necessary. Briefly describe materials and equipment you will use.
- **Data:** In tables (numbered & labeled). Include uncertainties and units.
- **Data Analysis:** Any and all needed calculations. Sample uncertainty calculations. Additional tables (comparisons, graphing, tables, slope/intercept etc.) and graphs.
- **Discussions:** Comparisons of data. Reasoning, discrepancies/agreements, commentary, sources of errors, classification of errors as systematic and random. Compare the obtained values to the expected ones (possibly using a test like the t-test). Comment on the agreement and on possible sources of errors (both random and systematic). Make sure the errors are reasonable given your results.
- **Conclusions:** (Include final values) Appropriate commentary on the agreement/disagreement of derived values. Report your values again and comment on their agreement with accepted values
- **References:**

Definition .2.1. A Hypothesis is a supposition or proposed explanation made on the basis of limited evidence as a starting point for the future investigation.

.3.0 Presentation Tips

The fundamentals of a good talk are a

- Good story: Clear beginning (describe the BIG PICTURE), clear plan (describe your topics and sub-topics if any), and clear answers (present the results and supporting evidence)
- Good slides: The three T's - Tell them what you are going to tell them (have a title page and outline slide), tell them (present the main slides), tell them what you told them (have a summary slide)
- Good presence: Dress for success (avoid wearing distracting items), engage the audience (look and talk to the audience), don't read the slides (and try to avoid reading from notes), body positioning and body language, and talk WITH not to.

In a normal talk you would usually budget 1 slide per minute, maximum. Text should be at least 22 pt, and you should avoid full sentences or lots of text on any given slide. Images should convey at most one or two concepts. Use reasonable colour contrast. Don't use unnecessary flashiness. KEEP IT SIMPLE SILLY.

Good powerpoint design:

- Keep it simple/consistent
- Use meaningful visuals
- Use predetermined colors/templates
- Don't overcomplicate/confuse
- Be distinctive or be related

When building a slide deck, have most of the work done already and let the template do the heavy lifting. Take your audience on a visual journey - go bolder on visuals and let the audience focus on you, not text. Know your key-takeaways - leave the audience something to take-away and make it obvious.

Repetition is a powerful and often useful tool in longer talks/presentations to remind audiences of your key takeaways. Progress bars or numbers can be powerful visual tools for guiding the audience. Use space appropriately, and leave suitable white-space between visuals/text. Less can often be more - use line animation when suitable to help guide visually. Make sure to establish a clear heirarchy in slides to guide the audience's eyes - animation can be a useful guide. Try to align visuals appropriately with suitable white spacing between columns.

Know your audience - what is their knowledge level? what level of professionalism do they expect?

.4.0 Paper Tips

Basic notes/points:

- Remember to reference the theory section properly
- The literature review is thorough - we compare with multiple peer-reviewed papers.
- Explain what every variable in an equation means
- Every statement that is not your own stated in the theory should be supported by peer-reviewed papers
- All equations must be referenced, and definitions should be referenced
- If steps aren't shown, the result must be referenced.
- Reference textbooks for definitions of basic terms.
- It is recommended to have 3-5 references per page of the report.

.4.1 Writing a Research Proposal

What is it? It is a document intended to convince others that you have interesting and important research. Why do you need it? To convince the reader of the value of your project/idea, and your competence. It is also to give yourself an opportunity to think through your research project, refine your ideas, and predict any challenges that may arise.

How to write one: (follow the instructions if given) i break down your proposal into point form. Choose your headings/subheadings. Lay out in point form what you will discuss. Use appropriate clean plain language.

.4.2 References

Follow the *Journal of Applied Physics* guidelines, <https://aip.scitation.org/jap/authors/manuscript>, <https://publishing.aip.org/authors/preparing-your-manuscript>, <https://aip.scitation.org/journal/jap>, https://publishing.aip.org/wp-content/uploads/2021/03/AIP_Style_4thed.pdf, and <https://paperpile.com/s/journal-of-applied-physics>. It is often best to use alphabetically listed references. For miscellaneous items that don't always appear it may be a good idea to see how APA does it before returning to the JAP guidelines. Choose how the numbering looks to be consistent.

Bibliographies

For an annotated bibliography, look at the *Purdue* and *Cornell* guides for how to build and annotate the bibliographies. Annotations are for you. Mention if its useful to you and if it

is trustworthy. You cite yourself like any other source (in square brackets put that it is a lab manual and unpublished).

An annotated bibliography is a list of citations, where each citation is followed by an annotation - a brief (100-150 words) descriptive and evaluative paragraph.

.4.3 Example-Outline

A large centered or left aligned title followed by names of the authors with footnotes referencing their affiliations. Date the paper, and include a centered clear abstract (at most 10 sentences), with keywords after the abstract. An experimental paper should include:

- **Introduction:**
- **Methods:**
- **Results and Discussions:**
- **Conclusions:**

Methods may be replaced with **Theory** if a non-experimental paper is being designed.

.4.4 Peer Reviewing

Helpful feedback follows the four following principles:

- **Goal Referenced:** Is linked to the goal specified by the receiver. Improving/developing presentation/writing skills.
- **Transparent:** Making the “how, what and why” behind editorial decisions clear to everyone - if possible/provided include page/line numbers for specific feedback. Content of the review is known not the identity of the reviewer.
- **Actionable:** Comment on things that actually can be improved. Provide feedback that strengthens the work.
- **User-Friendly:** Be objective. Back up each statement with an example. Be honest.