Table of Contents

Outline and Requirements of Engineering Physics 253

1. Objectives

The main objectives of this course are to gain experience with:

1. measuring physical quantities using standard instruments
2. keeping an organized record of your work in a laboratory notebook
3. determining experimental uncertainties and producing a credible assessment of your results
4. writing up experiments in the format of a scientific paper or technical report
5. writing a report in a short time frame. (This gives you experience in writing reports under pressure, a skill that is valuable in industry.)

These laboratory sessions are complementary to the lecture courses that you are taking, and provide an opportunity for you to develop skills that the latter do not address. Hopefully the experiments will be interesting in their own right.

2. Organization of the Laboratory

During this course you will perform 8 experiments. The first experiment *Introduction to Electronic Instruments* will be done in week 1and 2. In the 3rd week everyone will do lab 2 Electrical Impedance. In weeks 4-6 and 8-10 students will rotate through 6 other experiments according to a schedule that will be posted on the course website in week 3. Every other lab you will get your lab book marked at the end of the period. The analysis and sections of a written report will be handed in at the beginning of the next lab period (<u>**these are both done individually**</u>).

For the last lab done you will produce a full written formal report to be handed in week 12.

You have one lab period (3 hours) to complete the experiment you are assigned. If you look over the apparatus during laboratory time the week before, you will then have a picture of the apparatus in your mind as you study the instructions in preparation for the laboratory. Photographs of the laboratory equipment can be found on the course website if you need to refresh your memory as you read the instructions. **If you do not study the instructions before the laboratory period, you may not be able to complete the experimental work in the allotted time.** In exceptional circumstances it may be possible to return to the lab later to

complete an unfinished experiment but help from an instructor might not be available; see the faculty member in charge for permission if this proves to be necessary.
Caution: The instructions for each of the experiments provide an outline of the experiment - they are not reports or records of experiments and they should not be used as models for Formal Reports or notebook entries.

3. The Laboratory Notebook

The model to keep in mind when in the lab is the following. Assume that you are working on a project which is to be continued by another group when you leave. They will read your record of the afternoon's work, but you will not be able to talk to them. Thus a clear account of what you have done is essential, as well as some analysis to check that things are going well. *So, keep a logical, organized and chronological record of all of your work i.e. in the order of execution.*

You should always record data in the form shown on the instrument. Converting data as you write it down can lead to mistakes and leaves you without the original information if things turn out strangely (off by a factor of 4 or 1000). You will likely convert data to standard units eventually (e.g. you will probably measure and record current in mA but will eventually convert to A for calculations) but this can be done easily as part of your calculations in Excel, Matlab or Python.

The laboratory record must contain the following components:

1. The name of the experiment and the date you performed it at the start of each experimental record.

2. Sketches of the apparatus – use standard symbols for standard items such as voltmeters, power supplies, etc. (i.e. do not draw boxes with knobs and buttons - sometimes the instruction sheets do sketch the controls on the boxes but this is to help you identify and use them). Include model numbers of devices where available so that the experiment could be reproduced.

3. Chronological record of all procedures and measurements. Describe what you actually did; do not write 'procedure is the same as in the manual'. For each portion of the experiment start with a title or brief explanation of the purpose of that part of the experiment. Note any problems and how they were overcome, if they were. Record measurements and observations using tabular form when possible. Briefly explain how the data in each table were obtained, what is measured and how it is measured (e.g. are lengths measured with a wooden rule, steel rule, precision rule, vernier, micrometer, etc.) Estimate sensible uncertainties for all your data (see also next item). Remember, all tables should have headings and graphs should have captions.

5. When deciding on the accuracy of your data the first thing to consider is the precision of the instrument. This is the smallest increment you can reliably resolve on the scale. Then consider the reproducibility of the readings which may give a more realistic guide to uncertainty than the precision of the instrument. This is best done by each member making independent measurements of the same quantity. Finally, do you have reason to doubt the absolute accuracy

of the instrument you are using? If so, can you check it against another? Perhaps the calibration is off. Such checks will help reduce systematic errors.

6. Some preliminary calculations and/or graphs are done while the lab is in progress to ensure that things are proceeding as expected. Rough calculations and graphs can assist in detecting errors in procedure and omissions in the data while the equipment is still available and deficits can be corrected.

At the end of the lab have your TA do a quick check of your lab book to make sure there are no major omissions that need to be corrected before you leave. You will submit the lab book to the onQ drop box by the end of the day for marking.

4. Weekly "Analysis Appendix" and "Results and Analysis" Report Section *done individually*

Analysis Appendix

After the lab, analyze the data, producing appropriate tables and graphs, and answer any questions posed in the manual. The analysis should be written out in the order that you do it, including algebraic equations and sample calculations referring to data in the raw data tables as well as regression output associated with graphs and should be suitable for inclusion in the appendix of a report. Graphs are usually part of the analysis but should be kept on separate pages as these would eventually go in the main body of a report since they represent a polished summary of the data. Making graphs that fill an entire page also makes them big enough that trends in the data can be assessed more easily. See an instructor if you have problems with the analysis.

Tables:

- The reader should be able to look at a table and know what is in it without knowing the lab. The title should give the reader a clear idea of what was being tested, the conditions under which the data was collected and the purpose (be brief). For example, a data table that simply says "Current and Voltage readings" would be obscure. The reader would be left to wonder what the device under test was, what the conditions were (e.g. frequency of the source) and what the experimenter was trying to find out. Titles should be as brief and descriptive as possible.
- Extra information: if you can't include the information you need in the title briefly you can put it as a note below the table.
- Column headings should include title, units in parentheses and uncertainty if it is constant for the variable in the column.

Results and Analysis report section

The Results and Analysis section of a formal report summarizes what you did and found and will not include sample calculations or raw data tables. Sample calculations and raw data tables are left in the appendix where a reader can view them if they want to check anything but they shouldn't clutter up the main body of the report.

If you created a graph as part of your analysis you should indicate in the text what you graphed and why, refer to the graph by Figure number, and include the graph in the section. You should also describe key features of the graph and explain their significance.
Graphs:

- Titles should be descriptive as indicated for tables. There is no need to repeat information given in axis labels. e.g. *Amplitude attenuation for a low pass filter with R=50K*Ω *and C=0.5nF* rather than *V vs. log(ω)*.
- Axes should be labelled and have units in parentheses e.g. *Mass Flow Rate (kg/s)*.
- Trailing zeros on axis numbers should be trimmed to avoid unnecessary clutter e.g. 10 not 10.000.
- Regression lines should include equations with appropriate variables (not x and y if you are plotting V vs. I)
- Make graphs full page so the content can be assessed.
- Include error bars.
- If you are assessing the linearity include a residual plot.
- If there is more than one data set make sure the symbols can be differentiated easily.
- If you are including a theoretical comparison show the experimental data as points and the theoretical data as a curve without points.

Description of analysis should indicate what equations you used referring to them by number – the equations themselves are introduced in the Theory section of the report but if you aren't writing a theory section that week include the numbered equations in the appendix. Include theoretical values with uncertainty, or accepted values for comparison with your final results. Provide references for all accepted values.

5. Weekly report section

Starting with lab 2 an additional report section or two will be handed in starting with Abstract for the Impedance lab done in week 3 and then working through the sections in order: Introduction + Theory, Apparatus and Experimental Procedure, Discussion + Conclusions. Instructions on witting formal reports can be found in section 9.

6. The Formal Examination

During the term you will have had a chance to practice writing up each of the sections of a formal report and will have received feedback on your writing. For the last lab you do you will write a full formal report which you will hand in the last week of term. The marking sheet for the formal can be found on the course website, (see Formal Feedback).

7. Marking Scheme

The final mark for the course will be made up as follows:

- 6% for the lab book marks. (marked out of 10 every other week)
- 60% for the weekly analysis. (marked out of 10)
- 17% weekly report sections (each marked out of 10)
- 17% for analysis and Formal report of last lab. (marked out of 10)

Absence from Laboratory.

There might be cases where a student is ill and cannot attend the laboratory session. Such cases should be brought to the attention of the faculty member in charge so that arrangements can be made to make up the lab.

In all other cases a student needs authorization in advance to miss a laboratory.

8. The Formal Report

A formal report describes the results of a scientific investigation. Upon reading the report, the reader should be able to understand

- what you have done,
- why you have done it,
- how you have done it and
- what you have concluded.

The report should be self-contained, clear and concise. It should read smoothly, along the lines of an essay, and must be logically developed. The reader should be able to understand each section based on the previous sections only. The exception to this is the conclusions which should be able to be read independently. Use your own words – do not copy from the lab manual since it does not necessarily reflect what really happened and it is aimed at preparing students for the lab not reporting results to an outside reader. Avoid colloquialisms, ambiguities, clumsy phrasing, and overuse of particular words.

The report is divided into distinct sections: Abstract, Introduction, Theory, Apparatus and Experimental Procedure (or Method), Results and Analysis, Discussion, and Conclusions. Each of these will now be described.

Abstract

The abstract is a miniature version of the whole report. Often, people will not read beyond the abstract, so it is important that you are precise and specific. Your abstract should include the purpose, method, results and significance of the results. Often each of these can be stated in a single sentence or even less. The purpose and method can sometimes be combined, for example, "*The position of the centre of percussion of a rigid body supported from a hinge was determined by measuring the impulse produced on the hinge from the impact of a steel ball striking the body at various points*".

If you are measuring a particular quantity for which you have a theoretical estimate, then give both numbers (experiment and theory) with errors. If the quantity being measured has an accepted experimental value in the literature, quote this for comparison with your value, for example, "*The acceleration due to gravity was determined to be g=9.71+/- 0.08 m/s*^{2}*. This result agrees within 2 standard errors of the generally accepted value of g=9.81 m/s*^{2}."
Make sure that the abstract is self contained. This will require you to make decisions as to whether you need to add detail or remove detail so that you don't half say things.
The Abstract should be no more than a paragraph in length.

Introduction

This section sets the scene. It provides the reader with the background to understand the work in context and explains why the reader should be interested in this experiment. The introduction concludes with a statement of the objectives of the work such as, "*It is the purpose of this experiment to determine a value for the coefficient of viscosity of water by measuring the time for a fixed volume of water to flow through a narrow tube*."

Theory

The theory section gives the reader the background information necessary to understand the design of the experiment and the analysis of the results. Make sure to explain any assumptions made in the derivations of the equations, for example, air resistance was neglected, thermal losses are negligible etc. In some cases part of the procedure will be aimed at checking to see if these assumptions are reasonable.

Mathematical expressions should be written in sentences and variables defined when they are first used, for example, "*The magnetic field B at the centre of a plane circular coil of radius r with current I is given by*

$$B = \frac{\mu_0 n I}{2r}, \tag{1}$$

while the field along the x axis is given by

$$B = \frac{\mu_0 n I}{2(x^2 + r^2)^{3/2}}, \tag{2}$$

where n is the number of turns in the coil."

At this point the reader should have gained a clear understanding of the current state of knowledge as it applies to your experiment, and is now ready to read about how you actually did the experiment.

Apparatus and Experimental Procedure

This section should contain sufficient detail so that another researcher in the field could use your description to replicate the experiment. Be complete, accurate and precise. Do not copy the instructions given in the lab manual – you are not telling the reader what to do. Use your own wording to say what you really did and what actually happened (use **past tense** to describe what you did). Standard laboratory equipment does not need to be described, but if you used a special piece of apparatus to measure something, it is important to tell the reader.

A clear, labeled diagram of the apparatus should be included and referred to, for example, "*Using the apparatus shown in Figure 1, the time required for a ball bearing to fall was measured over the range of heights 20 cm to 150 cm.*" Neatness and clarity are important and good, legible labeling assists enormously in understanding the experiment.

Explain the main sources of possible error and estimate the magnitude of these errors.

Results and Analysis

This section presents your data and describes how you arrived at a final answer.

Usually a graph is the best way of presenting your results because it shows what the data looks like. (See the section on graphs in the Weekly Analysis section for information on graph format.) Number each graph as a figure in the order that they are presented in the text (diagrams are also figures and so graphs might not start at 1). Introduce each graph by discussing what is plotted and then point out its most important feature, for example, "*Figure 2 shows the temperature of the midpoint of the copper bar as a function of time. We see that the temperature becomes constant after approximately 90 minutes.*"

Explain how the data were analyzed and very briefly outline your error analysis. For example, "*Equation (5) indicates that a plot of x_m versus m should give a straight line of slope $\lambda/2$. A graph of x_m vs m for a frequency of 1000Hz is shown in Figure 3. The lack of a pattern in the distribution of points in the residual plot shown in Figure 4 indicates that the linear fit is appropriate. Using the slope determined by a least means square fit, λ was calculated to be 0.343 +/- 0.002 m .*" or "*The moment of inertia was calculated to be 5.1 +/- 0. 2 kg m^2 using Equation (6).*" Sample calculations and error analysis calculations should be put in an appendix where a reader can look at them if they choose to.

Supply theoretical values with uncertainty, or accepted values, so they can be used for comparison. Reference accepted values using a standard referencing format.

It is essential that you present your results carefully. Only then can you effectively discuss your results and present your conclusions.

Discussion

The Discussion describes the relation between your results and the theory. You must provide evidence that you have thought carefully about the meaning of the results.

Focus your attention on questions like these:

- Did the results turn out as expected qualitatively (e.g. trends, shapes of graphs, intercept of graphs)? If not can you explain why? e.g. "*We note that the straight line does not pass through the origin. This may be due to ...*" Make sure that the mechanism suggested is plausible.
- Quantitatively how do your experimental results compare to theoretical or accepted values? Can discrepancies be reasonably attributed to assumptions made in the theory (e.g. neglecting air resistance may reasonable for a lead ball, however, may not be if the ball is moving fast enough). If part of the procedure was aimed at checking the validity of an assumption in the theory make sure to discuss the implication of what you found here. If the discrepancies can not be explained, have you discovered anything new or unexpected?
- What are the strengths and weaknesses of your experimental design?
- Can you identify the experimental quantity which had the most effect on the final uncertainty and suggest a way to reduce it?

The reader should come away from the discussion with a clear message about what you have done and what you have discovered.

Conclusions

Provide an overview of the experiment and summarize your results and conclusions. You should also make suggestions for improvements (future work), however, complaints about faulty equipment or the amount of time spent are not appropriate.

References

Make sure you give a reference for all accepted values. Each reference that you cite should be listed at the end of the report. It should be as complete as possible so that the reader should have no trouble locating it.

Appendices

An appendix is used to put additional information about the experiment, where it will not interrupt the flow of the main text. The derivation of a complicated equation might go here, for example. You would refer to it in the body of the report with a sentence like "A full derivation of Equation (7) is given in Appendix 1". This way, the reader should be able to follow what you have done without reading the appendix but the details are there if the reader chooses to look at it.

In this course you will also include material so that the marker can check your work including data tables, sample calculations and error analysis. This material should be prepared ahead of time so that you can simply move it from your analysis binder to the appendix of your formal. Minimize references to this appendix as constant references disrupt the flow of the text. e.g. *All sample calculations and raw data tables can be found in Appendix n.*

A point form list of what we are looking for in formal reports can be found in the introductory material on the website under "Formal Feedback".

A sample formal report can be found at the end of the manual.

Introduction to Experimental Uncertainties

1. Introduction

You will study error analysis in some detail in the accompanying lectures. Here we give a brief summary of the main results which will enable you to begin making error analyses in an appropriate and acceptable manner.

We consider an *error* to be the difference between the calculated or measured value of a quantity and the *true* value. Notice that before performing the experiment the true value is unknown. If the true value was known then the estimation of the error would reduce to a trivial subtraction. Of course, if it were known, there would also be no motivation for performing an experiment. In most practical situations we must resort to error analysis to obtain the *uncertainty* in our value. This is an estimate of the expected error between our value and the true value.

How close your measurements come to the true value is determined by how carefully you control the various sources of error in the experiment. The size of an uncertainty conveys important information about the results. If you quote $t_1 = 10.00 \pm 0.01$ s and $t_2 = 10.0 \pm 0.1$ s then you are telling us that you put more confidence in the estimate of t_1 than t_2. These notes are primarily concerned with how we calculate the uncertainty. A careful analysis of the errors in an experiment may also focus attention on the part that most needs improving. Look out for this in your analysis of the experiments in the 253 laboratory.

2. Systematic and Random errors

The process of taking a measurement several times usually generates a spread in the measured parameter. These fluctuations are called *random errors*. A random error varies about the mean value of the quantity being measured and, if you take more and more data, you can make use of statistics to decrease the uncertainty in the final result. If this is the only type of error in an experiment, then in principle there is no limit to this decrease in the uncertainty and your result should converge to the true value given more and more data.

However, there is another class of experimental error called *systematic error*. Systematic errors may be generated by faulty calibration or poorly designed or constructed apparatus. Maybe your timer runs faster at higher temperatures. A dial might have backlash so you obtain a different result if you approach the reading from opposite directions. Or a screw might have been poorly cut so that its motion along the screw axis is not a linear function of its rotation as you assumed it to be. In some cases a systematic error might be constant throughout a set of measurements and, in principle, the data can be corrected later once the source of the error has been identified. This could happen if your voltmeter always reads 1% high. In any event, notice that, unlike the case of random error, repeating the measurements with the same apparatus does not eliminate systematic error, and the final result will generally not converge to the true value as you average more and more data.

> *Repeated measurements with the same apparatus neither reveal nor do they eliminate a systematic error. For this reason systematic errors are potentially more dangerous than random errors…G L Squires*

In summary, with random errors repeated measurements will give variable results that are equally likely to be higher or lower than the true value. With a systematic error, repeated measurements will give systematically higher or lower values than the true value.

3. Accuracy and precision

The *accuracy* of an experiment is a measure of how close the result comes to the true value. The *precision* of an experiment determines how precisely the result is determined without reference to the accuracy, i.e. how closely do successive determinations agree with each other? Words that have a similar meaning to precision are *resolution* and *repeatability*. Do successive results agree to 1 part in 100, or 1 part in 10,000, etc.?

These two parameters are not necessarily related. It is possible to have a very precise experiment that produces results reproducible to 0.1%, but is inaccurate in that the results always differ from the true value by 1%; this would presumably be due to a systematic error in the apparatus. Similarly it is possible to have an accurate, but imprecise, estimate of the true value. In this case repeated measurements might vary by 1% (i.e. the precision is 1%), but on averaging many results the answer might agree with the true value to 0.1%. This will happen only if there are no systematic errors in the apparatus. **A good experiment will simultaneously produce an accurate and precise estimation of the true value.**

4. Determining measurement uncertainties from scales

For determining uncertainty in measurements with some sort of scale (a ruler, a digital scale, a needle on a dial, etc.) the general rule of thumb is that the error in the measurement is ½ the smallest scale division, however, this is only applicable under ideal circumstances. In some cases the uncertainty is greater, for example if you can't get a ruler right up against what you are trying to measure, if the object has ill-defined edges, or if the digits on a digital display are fluctuating. Often, we can get the best idea of whether our uncertainty is reasonable by repeating the measurement, and we should always do so if we can. If possible have more than one person make the measurement and then look at the range of values. If there are a small number of repeated measurements (2 or 3) half the range of the values should be used as the uncertainty, this also works in the case of a fluctuating digital display. If a larger number of repetitions is made then the standard deviation of the mean can be used as the uncertainty. If repeated measurements are impractical you should consider how confident you are in your ability to read the scale, do you feel that you can read the scale consistently to half the smallest digit, if not to what level of precision?

5. Propagation of errors

(a) Functions of one variable

Let x be a measured quantity and y the required quantity defined by $y = f(x)$. This result is valid for both individual measurements and for the mean values. We shall assume that we are dealing with the mean values in what follows. The measurement of x will have a probable error Δx ($= \sigma_{\bar{x}}$ as defined above). Since

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \frac{dy}{dx}$$

the resulting error in y is approximately $\Delta y = (dy/dx)\Delta x$. For example, the area of a circle is given by $A = \pi r^2$. If the radius is measured to be $r \pm \Delta r$ then as estimation of the area will contain a propagated error ΔA given by

$$\Delta A = \frac{dA}{dr}\Delta r = 2\pi r \Delta r \quad \text{and a fractional error of} \quad \frac{\Delta A}{A} = 2\frac{\Delta r}{r}.$$

(b) Functions of several variables

This can be generalized to the case where you must measure several quantities a, b, …to obtain your final value of a parameter y, i.e. $y = f(a, b....)$. In this case we use partial derivatives as follows:

$$\Delta y = \frac{\partial f}{\partial a}\Delta a + \frac{\partial f}{\partial b}\Delta b \,...$$

However, if the variables are independent and random, then there is a good chance of partial cancelations of errors (some will be overestimates and some underestimates) and the best estimate of Δy is found to be given by

$$(\Delta y)^2 = \left(\frac{\partial f}{\partial a}\Delta a\right)^2 + \left(\frac{\partial f}{\partial b}\Delta b\right)^2 ...$$

The random errors are said to be added in quadrature when using such relations.
As examples, given a and b as defined, two particular results for the error Δy in y is as follows:

(i) $y = \pm a \pm b$ gives

$$\Delta y = \sqrt{(\Delta a)^2 + (\Delta b)^2}$$

(ii) $y = a^n b^m$ and also $y = a^n / b^m$ gives

$$\frac{\Delta y}{y} = \sqrt{\left(\frac{n\Delta a}{a}\right)^2 + \left(\frac{m\Delta b}{b}\right)^2}$$

For example, for a cylinder, $V = \pi r^2 L$ so $\Delta V/V = \sqrt{(2\Delta r/r)^2 + (\Delta L/L)^2}$

6. Sensitivity analysis

After you have completed an experiment, you may be looking to improve your procedure, possibly by looking at which variable caused the largest error. This can be done though a sensitivity analysis, which looks at each of the variables in the experiment and determines how much each contributes to the total error.

The formula to calculate the contribution from a variable is related to the equation for adding errors in quadrature. To find the factional error contribution of parameter a, then we take the first term in the equation, $\left(\frac{\partial f}{\partial a}\Delta a\right)^2$, and divide by the total error $(\Delta y)^2$. Multiplying this by 100 gives the value as a percentage as follows:

$$\%\ contribution\ from\ a = \left(\frac{\frac{\partial f}{\partial a}\Delta a}{\Delta y}\right)^2 \times 100\%$$

The higher the percentage the more that variable contributed to the total error. The results of the sensitivity analysis should be displayed as a bar chart showing the percent contribution of each variable.

7. Reporting measurements, significant figures

Uncertainties are usually quoted to 1 significant digit and NEVER more than 2 significant digits. Generally, if the first significant digit in the uncertainty is small (say, less than 5) and the second is in the midrange (say 4, 5, 6), then it may make sense to include a second digit since that digit results in an "appreciable change" in the uncertainty. For example, if an uncertainty is determined to be 0.15, then it might make sense to quote it as 0.15 instead of 0.2 since the additional decimal changes the uncertainty by 25% (an arguably appreciable change). Conversely, if the uncertainty is determined to be 0.92, then it might as well be reported as 0.9 since the 0.02 is a small relative change in the uncertainty.

When reporting results the least significant digit in the quoted value of the measured quantity is determined by the uncertainty, e.g. if the uncertainty is 0.02 then the result is reported to the nearest hundredth so a value calculated as 2.1386 would be reported as 2.14 $\pm$0.02. Also, it is much clearer to report the value and the uncertainty using the same exponents so write a result like $2.360\times10^{-3} \pm 3\times10^{-6}$ in the form $(2.360 \pm 0.003)\times10^{-3}$.

8. Straight Line Fitting to Data (Linear Regression)

Method of least squares

In an experiment you may have a set of data which you expect to follow a straight line and you wish to extract the slope and intercept. Although this can be done by eye, it is more conventional and less subjective to use a standard procedure. Methods have been developed that will give the best-fit slope and intercept together with an error estimate to indicate how reliable the parameters are. We will quote without proof the results for the case of a straight line fit to the data using a method based on the minimization of the squares of the differences of the fitted points from the straight line.

The straight line is taken to be of the form

$$y = A + Bx$$

and is to be fitted to a data set containing N pairs (x_i, y_i) where the error is assumed to be only in y_i, i.e. x_i is arbitrarily accurate. The best estimates of A and B are found to be

$$A = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum (x_i y_i)}{\Delta} \quad \text{and } B = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\Delta}$$

where $\Delta = N \sum x_i^2 - \left(\sum x_i\right)^2$. All sums are from $i = 1$ to N. Often we wish to force the fitted line to go through the origin, i.e. it might be physically incorrect to have an intercept. In this case we have

$$A = 0 \qquad \text{and} \qquad B = \frac{\sum (x_i y_i)}{\sum x_i^2}$$

After obtaining A and B, we calculate the deviations h_i of all the measured y_i from the fitted line, i.e. $h_i = y_i - (A + Bx_i)$. We will write the standard deviation of the distribution of the h values as σ_y; this is defined by

$$\sigma_y = \sqrt{\frac{\sum h_i^2}{N-2}}$$

The *standard error* estimates σ_A and σ_B in A and B respectively are then given by

$$\sigma_A = \sigma_y \sqrt{\frac{\sum x_i^2}{\Delta}} \text{ and } \sigma_B = \sigma_y \sqrt{\frac{N}{\Delta}}$$

When the line is forced though the origin, the factor $N-2$ is replaced by $N-1$ and the error in the slope is

$$\sigma_B = \sigma_y \sqrt{\frac{1}{\sum x_i^2}}$$

The results of a least-squares fit are quoted in the form $A \pm \sigma_A$ and $B \pm \sigma_B$. The errors have the normal meaning of a probable error in that there is a 68% probability of finding the true result within the quoted range.

N.B. Straight line fitting is done by spreadsheets such as Excel (select *Data Analysis/Regression)* **and other software and all use the method just described.**

Weighted regression

If the uncertainties vary significantly between points <u>and</u> you can measure the uncertainties σ_i accurately then a weighted fit based on the uncertainty in each point is better than the unweighted fit just described. The rule of thumb for “significantly" is around a factor of 4.

Using weights $w_i = 1/\sigma_i^2$ best estimates of A and B are found to be

$$A = \frac{\Sigma w x^2 \, \Sigma w y - \Sigma w x \, \Sigma w x y}{\Delta}$$

and

$$A = \frac{\Sigma w \, \Sigma w y - \Sigma w x \, \Sigma w y}{\Delta}$$

where now

$$\Delta = \Sigma w \Sigma w x^2 - (\Sigma w x)^2$$

The *standard error* estimates σ_A and σ_B in *A* and *B* respectively are then given by

$$\sigma_A = \sqrt{\frac{\Sigma w x^2}{\Delta}} \text{ and } \sigma_B = \sqrt{\frac{\Sigma w}{\Delta}}$$

This algorithm calculates the error in the slope using the measured errors through the inclusion of the weights rather than the variation of the points from the predicted line. If the uncertainties in the measurements are under or overestimated then so will the error in the slope.

Residuals

In order to get an idea of how well our model "fits" the data, one can look at the "residuals" of the fit. For a data point, (x_i, y_i), with a model $y_{model}(x)$, the residuals, $R(x_i)$, are defined as:

$$R(x_i) \equiv y_i - y_{model}(x_i)$$

and correspond to the difference between the data and the model prediction. For a linear fit the residuals are equal to the deviations from the straight line that were calculated in the least mean square fit algorithm:

$$R(x_i) = h_i = y_i - (A + Bx_i)$$

Again, if we assume that the errors in the xi are negligible, then the error in the residuals are given by:

$$\sigma_{R(x_i)} = \sigma_{y_i}$$

That is, the error on the residuals is given by the error in the data points. If the model is a good representation of the data, then the residuals are expected to be symmetrically distributed about a mean of 0. If all of the residuals were positive or negative, then one would question whether the model is a good fit to the data. Similarly, if the residuals show a trend, then one would also suspect that there is an issue with the model.

9. The mean and standard deviation

This section and the remaining sections will focus on the characteristics and statistical properties of random errors. If repeated measurements of the same quantity x are made a distribution of the measured result will be generated provided the resolution of the instruments is sufficiently high. This distribution may be due to uncontrolled variation of conditions under which the measurement is made, difficulty in judging when conditions for the measurement are met, the fact that the measured quantity is an average of one which varies in space and time, or it might arise from inherent random fluctuations in the quantity being measured, e.g., due to thermal

noise or the randomness of nuclear decays. We call the set of values so obtained a *sample* and an example is shown in Fig. 1.
To analyse this further we require some statistical measures to characterize the distribution. The most frequently used statistical measures are the *mean* $\bar{x}$ and the *standard deviation* σ_x. If N measurements of a variable $x_1, x_2, \ldots, x_i, \ldots x_n$ are obtained, the mean, or average, is defined by

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i .$$

The standard deviation σ_x is a measure of the width, or spread, of the distribution and is given by

$$\sigma_x = \sqrt{\frac{1}{(N-1)}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

While the sample mean gives the average value of the measurement, the sample variance and standard deviation indicate the average distance between the measurements and the sample mean. That is, whether the measured values are all really close to the sample mean, or whether they are spread out. The sample standard deviation is also a representative measure of the uncertainty in a single measurement.

As the number of values N becomes larger and larger, and finally infinite, the sample distribution approaches what is called the *parent distribution*, yielding the parent mean X and the parent standard deviation σ. It is X that we are trying to measure.

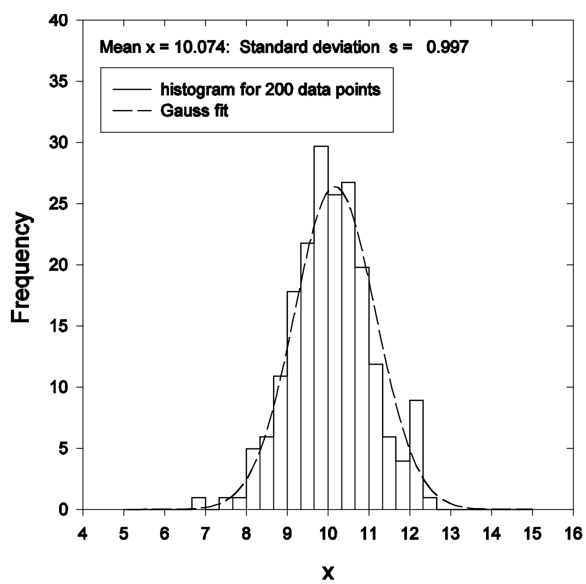


Fig. 1 This shows the result of taking 200 measurements of a quantity x. The number (or frequency) of the values falling into each 'bin' of width 1/3 are given as a function of x. (The sum of the numbers in all bins is thus 200). The parent mean is X=10.00 and the standard deviation is $\sigma = 1.00$. This sample has a mean of $\bar{x} = 10.074$ and standard deviation of σ_x=0.997. The dashed line is a Gaussian using $\bar{x}$ and σ_x as the parameters in Eq. (3).

10. Probable error of the mean

The sample gives a values of $\bar{x}$ and σ_x which are estimates of X and σ. If we took another sample (i.e. another N measurements of x_i) we would generally get different values of $\bar{x}$ and σ_x as estimates of X and σ. What then is the best measure of the uncertainty or error between $\bar{x}$ and X? At first sight you might think that we just quote $\bar{x} \pm \sigma_x$ *but this is NOT correct*. Clearly σ_x is a measure of the error between x_i and X for any single measurement, but we have N measurements so surely the difference between $\bar{x}$ and X is smaller than σ_x.

In fact the various values of $\bar{x}$ that we obtain from different samples have their own distribution. This distribution is *narrower* than that for the individual measurements x_i of a sample. To analyse this problem we need to know the details of the parent distribution for x. Typically it is assumed to be *Gaussian*. In this case the probability $P(x)$ of finding a value x in the range x to $x+dx$ is given by

$$P(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mathrm{X}}{\sigma}\right)^2\right]dx$$

This has a mean of X and standard deviation of σ. If $P(x)$ is integrated over all possible values of x, i.e., $-\infty$ to $+\infty$, then we obtain unity as expected for probabilities.

It is then found that the distribution for the various sample means $\bar{x}$ is also Gaussian. If $P(\bar{x})$ is the probability of finding a value of $\bar{x}$ in the range $\bar{x}$ to $\bar{x}+d\bar{x}$ then

$$P(\bar{x})d\bar{x} = \frac{1}{\sigma_\mu\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\bar{x}-\mathrm{X}}{\sigma_\mu}\right)^2\right]d\bar{x}$$

where $\sigma_\mu = \sigma/\sqrt{N}$. It is the width of this distribution, quantified by σ_μ, which is our best estimate of the uncertainty in in our measurements, i.e., the expected difference (the error) between X and $\bar{x}$ for any sample is σ_μ. Of course we do not have measurements for the parent distribution so we use our results for the sample distribution to give our best estimates of X, σ and σ_μ. These are

$$\mathrm{X} \approx \bar{x}, \qquad \sigma \approx \sigma_x, \quad \sigma_\mu \approx \frac{\sigma_x}{\sqrt{N}}$$

We call $\sigma_{\bar{x}} = \dfrac{\sigma_x}{\sqrt{N}}$, the *probable error of the mean* and so our best estimate of X and its uncertainty from our sample are

$$x_{best} = \bar{x} \pm \sigma_{\bar{x}}$$

As an example, 10 readings of a time interval produce the following values:

20.28, 21.26, 20.96, 20.70, 20.31, 21.16, 20.60, 20.36, 20.55, 19.95 sec

We find $\bar{x} = 20.61$, $\sigma_x = 0.42$, $\sigma_{\bar{x}} = 0.42/\sqrt{10} = 0.13$, so that $x_{best} = 20.61 \pm 0.13$ sec. It is also possible to show that the probability that $\bar{x}$ falls in the range $\mathrm{X} \pm \sigma_{\bar{x}}$ is 0.68. Likewise, between $\mathrm{X} \pm 2\sigma_{\bar{x}}$ it is 0.95, and between $\mathrm{X} \pm 3\sigma_{\bar{x}}$ it is 0.997. These are often called 68%, 95% and 99.7% *confidence limits*.

After completing our measurements and analysing the data, we usually compare our result with the *accepted* value of the quantity, if there is one. The accepted value is the best estimate of the true value that you can find in the literature. If you have considered all your sources of error, both systematic and random, your final value should agree with the accepted value within the uncertainties. When doing this, remember that the uncertainties quoted are typically for one standard deviation. If the results do not agree within 3 standard deviations, i.e. at the 99.7% confidence level, there is probably something badly wrong somewhere.

Finally, note that all these error estimates do not include any systematic errors in the apparatus. These must be considered independently.

11. Bibliography

The following books provide more information about error analysis at various levels of difficulty. There are many more.

1. *An Introduction to Error Analysis* by J. R. Taylor, University Science Books, (1997). An excellent introduction to error analysis.

3. *Data Reduction and Error Analysis for the Physical Sciences*, P R Bevington and D K Robinson, McGraw-Hill Book Company (2002). A comprehensive text on error analysis which goes beyond our present needs but will remain an invaluable reference.

Introduction to Electronic Instruments and Low Pass Filter

In this introductory experiment you will explore voltage dividers and a low pass filter while being introduced to oscilloscopes, signal generators, multimeters and breadboards.

1. Introduction

The main aim of this laboratory is to introduce you to basic electronic measuring instruments. The important functions and properties of these instruments are outlined in the following sections. This is followed by some basic experiments which illustrate the various features and also introduce you to the characteristics of resistance/capacitance (RC) filters. For some of the measurements you will build simple circuits on a *solderless breadboard*. These instruction sheets will also provide a reference for when you encounter instruments in the 253 and other labs. When you become more experienced you will find it an advantage to refer to the manufacturer's instruction sheets to get the most out of any instrument.

2. Multimeters

Most multimeters are digital (and so are often referred to as DMMs - digital multimeters). They are designed to measure voltage (DC and AC), current (DC and AC) and resistance (DC only). Sometimes the meter will also measure capacitance and/or temperature. The different functions might require that you use different input terminals so be careful when you connect the meter.

(a)Voltmeters

For measuring voltage, we use the multimeter as a *voltmeter* and connect the input terminals across (i.e., in parallel with) the voltage being measured. The symbol for a voltmeter is a capital V surrounded by a circle. Fig. 1 shows a voltmeter placed to measure the voltage v (i.e., potential difference) across the resistance R.

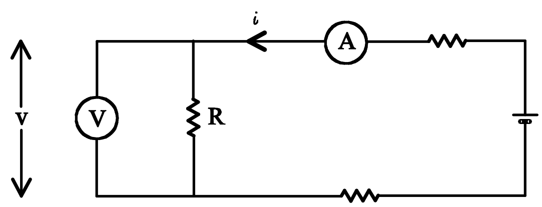


Fig. 1 Measuring voltage and current

Naturally, we want to make sure that connecting the meter has no significant effect on the operation of the circuit. For practically all purposes the voltmeter can be viewed as a simple resistance looking into its input terminals, i.e., its *equivalent circuit* is a resistance. This resistance is called the *input resistance*, say R_{in}^{V}, of the voltmeter. Because we want the voltmeter to have the minimum effect on the voltage being measured, then R_{in}^{V} should be as large as possible (ideally infinite). A typical voltmeter might have R_{in}^{V} in the range of 10 $M\ \Omega$ - 1 $G\ \Omega$ and the one supplied here has $R_{in}^{V} = 10\ \mathrm{M}\Omega$. As an example, in Fig. 1 we want $R_{in}^{V} >> R$ if connecting the voltmeter is to have a negligible effect. If R = 1.0 $M\ \Omega$ and $R_{in}^{V} = 10\ M\ \Omega$, then connecting the voltmeter across R_1 would give a parallel resistance of R and R_{in}^{V} of about 0.9 $M\ \Omega$ which is a 10% change and certainly not negligible.

(b) <u>Ammeters</u>

When measuring current, we use the multimeter as an *ammeter* and connect the input terminals in series with the current being measured. The symbol for an ammeter is a capital A surrounded by a circle. Fig. 1 shows an ammeter placed to measure the current through the resistor R. Here the ammeter is measuring the current through *both* R and the voltmeter. If $R_{in}^{V} >> R$ then any current through the voltmeter will be negligible. Again the equivalent circuit of the ammeter is a resistance, say R_{in}^{A}. We want the ammeter to have a minimum effect on the circuit so R_{in}^{A} should now be as small as possible (ideally zero). In Fig. 1 we want $R_{in}^{A} << R$ if the ammeter is to have no significant effect on the circuit.

R_{in}^{A} is usually not given explicitly by the manufacturer. Instead the ammeter specifications give the voltage that appears across the ammeter terminals when the current through them equals the maximum (i.e., full scale reading) on the particular range that the ammeter is set to. For example, the manufacturer might specify that there is a voltage drop of 0.10 V when the meter reads full scale on the 1.0 A range. Using Ohm's law this tells you that $\mathrm{R}_{in}^{A} = 0.10/1.0 = 0.10\Omega$. The voltage across the terminals for any other current i measured on this same range is then $R_{in}^{A} i = 0.10i$ Volts. Typically the meter is designed so that the voltage drop at full scale has the same value whatever the range. This means that R_{in}^{A} is different on each range. For the meter you are using here, the voltage drop is 0.10 V for any full scale current reading.

(c) <u>Measuring AC quantities</u>

When used to measure AC voltage or current, the meter gives the rms (root-mean-square) value of the voltage or current. Thus if the voltage being measured is $v(t)$, then the voltmeter reads

$$V = \sqrt{\frac{1}{T}\int_0^T v^2(t)\,dt} \quad (1)$$

where T is the period of the waveform. The limits of the integral are arbitrary as long as they span a full period. Thus if the voltage being measured is $v_\circ \cos(\omega t)$, where $\omega = 2\pi f$, then it reads $V = v_\circ/\sqrt{2}$. Meters give accurate values of the rms voltage only if the waveform is sinusoidal.

Similarly, when measuring an AC current $i(t)$ the ammeter reads the rms value of $i(t)$ given by

$$I = \sqrt{\frac{1}{T}\int_0^T i^2(t)\,dt} \quad (2)$$

Multimeters have a rather restricted range of frequency when measuring AC quantities. Your meter is accurate over the frequency range about 30 Hz to 10 kHz.

(d) Ohmmeters

When measuring resistance, the multimeter is used as an *ohmmeter*. In this mode it supplies its own DC current i (which depends on the scale being used) to the resistance under test and measures the voltage drop across the resistance v; the resistance is then v/i. The resistance must be isolated (disconnected) from any circuit before measuring.

3. Signal Generators

A signal generator is designed to give particular voltage signals at its output connector (often called output jack). In the present case the signals are periodic and are one of sinusoidal, triangular or square. In your generator the waveform is chosen by push-button switches. The frequency $f = (1/T)$ is continuously variable from ~ 0.5 Hz to ~ 5 MHz using push-button range switches and two knobs called COARSE and FINE. The frequency is displayed on the panel meter.

On your generator there are 2 output jacks labelled OUTPUT and TTL/CMOS (and an input jack labelled VCG INPUT which we will not use). All of these use a type of jack called a BNC connector designed for use with coaxial cables (Figure 2a). Note that the outside of these connectors are grounded and the signal is carried on the inner pin and connected to the centre wire of the coaxial cable.

Fig. 2a BNC Connector

The jack labelled OUTPUT is where the signal appears. Figure 2b gives the *equivalent circuit* of the signal generator. The generator behaves like a battery with an oscillating emf, which has an internal resistance of R_{out}. Note the symbol used for the sinusoidal signal generator (but R_{out} is typically not shown on circuit diagrams). Usually signal generators have an *output impedance* $R_{out} = 50\,\Omega$ but in some generators it might be $600\,\Omega$. The amplitude $v(t)$ of the output can be varied by the OUTPUT LEVEL knob and the push-button switch labelled -20db. When pressed, this button decreases the output amplitude by a factor of 10. There is also a push-button switch and a knob labelled DC OFFSET which add a variable ± dc voltage to the output when pressed.

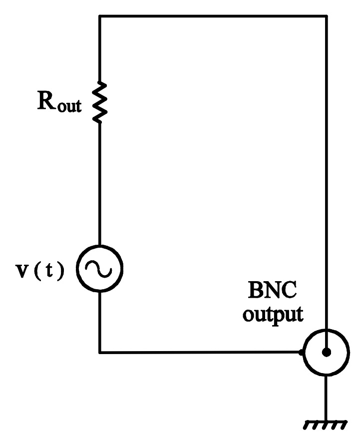


Fig. 2b Equivalent circuit of generator

The TTL/CMOS output jack gives a square wave which switches between zero and a positive voltage of a few volts. This has many uses, but one involves supplying synchronization pulses to an oscilloscope (see below).

4. Oscilloscopes

Oscilloscopes are designed to examine time varying voltage signals. The oscilloscope you will use here is useful only for repetitive signals, i.e., signals that repeat over and over, but you will meet others (e.g. in the experiment *Young's Modulus*) which will display a non-repetitive signal. There are two versions of oscilloscopes used in this laboratory; the controls are very similar in function though they are placed on different regions of the front panels and may also be labelled slightly differently.

Time Base

Usually the (horizontal) x axis is used to display time t. An internal circuit called the time base causes the bright spot on the screen (the location of the impact of the electron beam) to move across from left to right at a steady rate called the sweep rate. When the spot reaches the right hand side, it moves back to the left hand side essentially instantaneously and begins again (more about this below). The switch labelled TIME/DIV sets the sweep rate and the division referred to is 1 cm on the display. This switch also has an inset knob which allows a continuously variable sweep rate. Set this to the calibrated position if you want the values on the switch positions to be accurate. There is also a knob labelled by arrows pointing right and left which shifts the display horizontally so you can centre it on the screen if required.

Input Channels

You can display 2 independent signals on the (vertical) y axis. These are introduced via 2 input jacks, probably labelled CH1 or X and CH2 or Y, referred to as Channel 1 and Channel 2. You can use these input signals to produce two y,t plots simultaneously. An example of a simple circuit connected to an oscilloscope is given in Fig. 3. Note that you can consider the oscilloscope to be a voltmeter which displays the instantaneous voltage as a function of time.

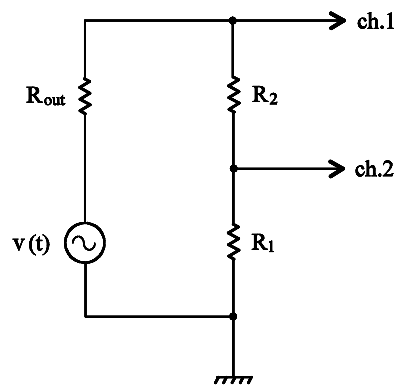


Fig. 3 A circuit with 2 signals going to an oscilloscope.

Just as with the voltmeter discussed above, there is an input resistance associated with each channel. This is typically 1 $M\ \Omega$, a value much lower than for your multimeter so that there is a greater likelihood that the oscilloscope will affect the circuit to which it is connected.
All the input jacks to the oscilloscope are BNC connectors which have their outer parts grounded, so each channel only measures voltages with respect to ground. If you are not clear about what this means, measure the resistance between the outer parts of different connectors (including those on the signal generator) with an ohmmeter. *It follows that the signal generator and input channel grounds must all be connected to the same point on the circuit.* Thus in Fig. 3 CH2 measures the voltage across R_1 and CH1 measures the voltage across $R_1 + R_2$.

The input signals are amplified so that a convenient display size can be produced on the screen. A selector switch enables the input signals to be AC or DC coupled to the amplifier. With DC selected, both AC and DC components of the signal are amplified equally. With AC, a capacitor is inserted in series with the signal which blocks the DC component so that only the AC components are amplified, typically down to a frequency of $\leq 1 \ Hz$.

The sensitivity of the amplifier is shown on the VOLTS/DIV switch. Again the division referred to is 1 cm on the screen. As with the TIME/DIV switch, the VOLTS/DIV switch also has a small inset knob which allows for a continuously variable sensitivity; set this to calibrated if you want the values on the switch positions to be accurate.

Because there is only one electron beam in the oscilloscope, you can only display one channel (signal) at a time. However, there is a switch labelled MODE which enables you to choose how to display the 2 channels. The options are CH1, CH2, ALT, CHOP and ADD. In the ALT position each channel is displayed alternately. Often the traces are displayed so fast that they seem to be simultaneous. In the CHOP position the display rapidly switches between the 2 channels as the beam moves across the screen so the 2 traces seem to be displayed at the same time. You cannot see the switching between channels when the TIME/DIV is rather long but at high sweep rates it becomes obvious.

Finally, there is a special position on the TIME/DIV switch labelled XY. At this position the signal into CH1 or X is used to drive the spot horizontally, and the signal into CH2 or Y drives it vertically so you can make an xy plot with the 2 signals.

Trigger

There is an input jack on the control panel labelled TRIG IN. To understand this feature, imagine applying a simple sinusoidal signal in Channel 1 and displaying a y,t plot (using the internal time base). The first time the spot moves across the screen it displays the sine wave starting at a random point (i.e., phase) on the wave. The spot then returns to the LHS and begins again at another random phase on the wave. As a result of the overlap of many such traces, we cannot see anything useful on the screen.

To prevent this happening, the sweep is arranged to always begin at the same phase of the wave. This can be done internally using the signal itself (with the INT or CH1 setting on the SOURCE switch) or by using the EXT setting and sending a repetitive external signal, which is synchronized with the sine wave, to the TRIG IN jack. We usually use the latter setup when such a *sync* signal is available. There is also a LINE setting, which triggers the scope at the 60 Hz mains frequency, and an AUTO setting which is basically free running but is often useful for seeing if you have any signal at all.

5. Experiment

Part 1: Voltage divider with two resistors

A voltage divider consists of two components in series that produce an output voltage that is a fraction of the input voltage. The simplest case consists of two resistors where the input voltage v_{in} is applied across the series combination of the resistors and the output v_{out} is the voltage across one of the resistors. Using Ohm's Law[1] for the example shown in Figure 4 where v_{in} is the voltage across both resistors (i.e. the voltage from the generator) and v_{out} is the voltage across R_1 we find that

$$v_{out} = v_{in}\frac{R_1}{R_1 + R_2} \qquad (3)$$

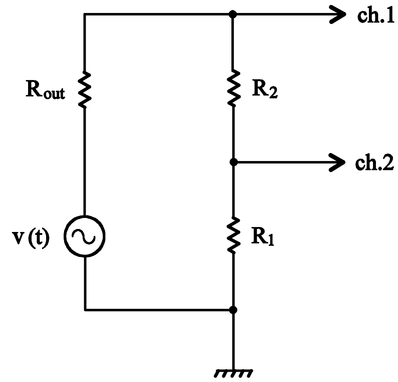


Fig. 4 Resistive voltage divider circuit. Recall that each channel measures voltages with respect to ground thus ch2 measures the voltage across R_1 which is v_{out} and ch1 across $R_1 + R_2$ which is v_{in})

Use your breadboard to build the voltage-divider circuit shown in Figure 4 using values in the ranges R_1 =1kΩ-2kΩ and R_2 = 50kΩ-100kΩ. Use your multimeter as an ohmmeter to measure the exact values of the 2 resistors you select.

Using the coaxial cable with a BNC connector on each end connect the signal generator to the breadboard as shown in Figure 5. Apply a sinusoidal voltage, v_{in}, from the generator to both the resistors in series.

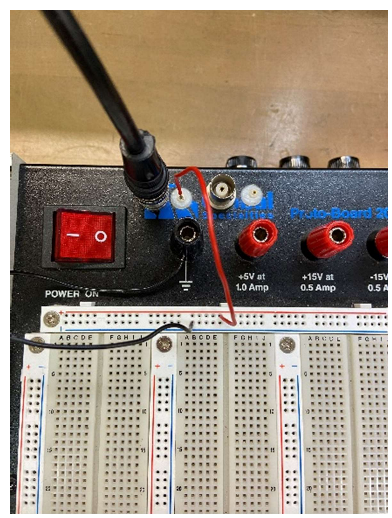
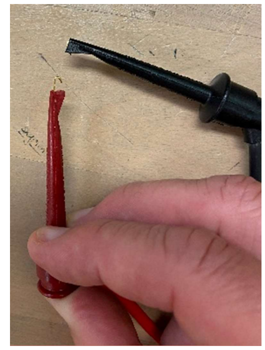


Fig. 5 The wire socket beside the BNC connector is used for making connections from the centre of the BNC socket to any point on the board.

[1] $v_{in} = i(R_1 + R_2)$ so $i = \frac{v_{in}}{R_1 + R_2}$ now if $v_{out} = v_{R1}$ then $v_{out} = iR_1 = \frac{v_{in}}{R_1 + R_2}R_1$

External triggering of the oscilloscope:
Connect the TTL/CMOS output of the signal generator to the TRIG IN jack on the oscilloscope and set the SOURCE switch to EXT. Adjust the LEVEL knob if necessary to stabilize the display. Now notice that you can change the amplitude and DC offset of the signal generator and the displayed signal is always stable. Leave the external trigger in place for the rest of the experiments.



Use the oscilloscope to determine the ratio of the voltage across R_1, v_{out}, to that from the generator, v_{in}, i.e., v_{out}/v_{in}, at a frequency of $\sim$ 5 kHz. Use the scope probes with the hooks to connect to the circuit elements. Does this agree with the expected value? Are the 2 signals in phase? Does the ratio and phase depend on frequency over the range 500 Hz to 500 kHz?

Now try using the multimeter set to AC to measure the voltage across the resistors. Do your measurements agree with the ones you obtained using the oscilloscope?

Part 2 Low Pass Filter

A low pass filter is a circuit designed to attenuate signals at frequencies above a specified cutoff frequency while passing signals below the cutoff frequency. Low pass filters are often used to reduce electrical noise in a signal. A simple low pass filter can be created by replacing one of the resistors in the voltage divider from part 1 with a capacitor (C). Using Ohm's Law with the impedances of the circuit elements gives a result similar to the previous case: $\mathbf{V_{out}} = \mathbf{V_{in}} \dfrac{\mathbf{Z_C}}{\mathbf{Z_C} + \mathbf{Z_R}}$

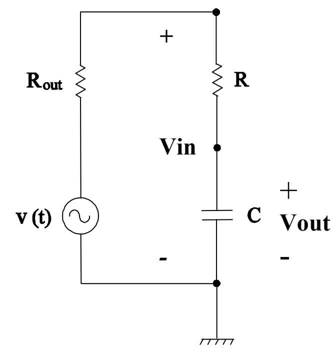


Figure 6: Low pass filter

where $\mathbf{V_{in}}$ and $\mathbf{V_{out}}$ are the phasors representing the input and output sinusoids respectively and $\boldsymbol{Z}$ is the impedance of the circuit element. Substituting $\boldsymbol{Z_C} = \dfrac{1}{j\omega C}$, and $\boldsymbol{Z_R} = R$, where $\omega = 2\pi f$ we get

$$\frac{\mathbf{V_{out}}}{\mathbf{V_{in}}} = \frac{\dfrac{1}{j\omega C}}{\dfrac{1}{j\omega C} + R} = \frac{1}{1 + j\omega R\, C} \tag{3}$$

Equation 3 is called the transfer function (***H***) and defines the relationship between the amplitude and phase of the signals at the input and output of the filter. Usually the transfer function is displayed by plotting two graphs, one showing the magnitude of the transfer function as a function of frequency and one showing the phase of the transfer function. The expression for the magnitude and phase of the transfer function are given in equations 4 and 5 respectively[2].

$$\left|\frac{\boldsymbol{V_{out}}}{\boldsymbol{V_{in}}}\right| = \frac{1}{\left(1 + (\omega R\, C)^2\right)^{1/2}} \tag{4}$$

$$\angle\left(\frac{\boldsymbol{V_{out}}}{\boldsymbol{V_{in}}}\right) = \tan^{-1}\left(\frac{0}{1}\right) - \tan^{-1}\left(\frac{\omega RC}{1}\right) = -\tan^{-1}(\omega R\, C) \tag{5}$$

---

[2] Recall :

a) for $\boldsymbol{C} = A + jB$ we can write $\boldsymbol{C} = M\angle\phi$ where the magnitude $M = \sqrt{A^2 + B^2}$ and the phase angle $\phi = \tan^{-1}\left(\dfrac{B}{A}\right)$

b) $\left|\dfrac{\boldsymbol{C_1}}{\boldsymbol{C_2}}\right| = \dfrac{M_1}{M_2}$ and $\angle\dfrac{\boldsymbol{C_1}}{\boldsymbol{C_2}} = \angle\boldsymbol{C_1} - \angle\boldsymbol{C_2}$

Once the transfer function is known, it is easy to predict the response of the circuit to a sinusoidal signal. If $v_{in} = v_0 \cos(\omega t)$ then $v_{out} = \dfrac{v_0 \cos(\omega t + \phi)}{\left(1 + (\omega RC)^2\right)^{1/2}}$ where $\phi = -\tan^{-1}(\omega RC)$.

Use your breadboard to build the circuit shown in Figure 6. Select components in the ranges R = 1 to 2 kΩ an C = 0.02 to 0.04 μF. Measure the resistance (R) and capacitance (C) using the multimeters found on either side of the storage boxes. The uncertainty in the resistance measurement is (0.5%)R +/- 1 in the last digit, and the uncertainty in the capacitance is (0.5%)C +/- 1 in the last digit. Record the value of R and C and the uncertainties. [3]

Connect two scope probes to your circuit so that you can display v_{in} and v_{out} on the oscilloscope at the same time. (You won't be able to clip the scope leads to the legs of the capacitor so you can insert a wire in an adjacent hole and clip to that.) The ratio of the magnitudes of v_{out} and v_{in} will give the magnitude of the transfer function while the time shift between the signals can be used to calculate the phase ϕ. Note: you will need to record the sign of the phase shift. See Appendix 1 for information about making these measurements.

Set the signal generator to display a sinusoidal signal at a frequency of approximately 100 Hz and adjust the amplitude to the maximum for the signal generator (~ 20 volts peak to peak). At 100 Hz the ratio of the amplitudes of the 2 signals should be close to unity and there should be a negligible phase difference between them. If the amplitudes are not the same, check for a systematic error in the calibration of the amplifiers by measuring both signals using the same scope probe. If there is still a difference, check to make sure the frequency of the signal and values of R and C are in the specified range. Before collecting data scan through frequencies from 100 Hz to 100 KHz and make sure the circuit behaves as expected i.e. low frequencies are passed unattenuated and high frequencies are attenuated. If this doesn't occur check your resistor and capacitor values with a multimeter. Once you are satisfied that the circuit is correct, measure the amplitude of v_{in} and v_{out} and the phase shift at the following frequencies in Hz: 100, 300, 1k, 3k, 10k, 30k and 100k. Note, there is no need to set *f* accurately; any values within $\pm 25\%$ of those given are fine so long as you record the exact frequencies used. When you analyze your results it is acceptable to ignore the uncertainty in the frequency. Adjust the TIME/DIV and VOLT/DIV knobs at each frequency as required. Note that a division is a box. Remember to record the time scale and voltage scale settings for each measurement and also record your estimate of the uncertainty in each measurement in term of a fraction of a division. The uncertainty will change when you change the voltage and time scales. Calculations of voltage and phase shift can be calculated later in whatever programming language you chose.
Here is an example table.

f (Hz)	Vin (div) +/- _ div	Vin scale	Vout (div) +/- _ div	Vout scale	Δt (div)	Time scale

[3] Example: Let's say you measured 1.021 kΩ. The last digit shown on the meter is in the thousandths so "1 in the last digit" would be 0.001kΩ. The uncertainty would be +/- (0.005*1.021 + 0.001) kΩ.

Analysis

(1) Plot the theoretical magnitude of the transfer function (equation 4) as a function of frequency over the same range as the experimental measurements. Use a log scale for the horizontal axis (frequency) and plot the magnitude of the transfer function in units of dB. For example, in MATLAB if you defined two arrays `H` and `f` that contained the samples of the magnitude of the transfer function and the corresponding frequencies respectively, the following command would produce the desired graph:

```
figure(1); semilogx(f,20*log10(H));
```

(2) On the same graph, plot two curves that define the range of uncertainties in the theoretical values due to the uncertainties in R and C. Here, the uncertainties in theoretical values are probably pretty small. However, you should do the calculation to verify this. A very useful trick for calculating the uncertainties is given in Appendix 3. If you define an array dH that contains the uncertainty in the theoretical values of H at each frequency, then the following MATLAB command would add dotted lines on either side of the theoretical curve that define the range of uncertainties:

```
hold on; plot(f,20*log10(H+dH),':',f,20*log10(H-dH),':') hold off;
```

The ‘hold on’ command allows you to plot more than one curve in the same figure window and the ‘hold off’ command releases the figure window for the next graph. Note, when you set up the frequency array, make sure you step the frequency in equal powers of frequency so that the points are evenly distributed along the frequency axis. For example, in MATLAB you could use the following commands to define the frequencies used in your plot:

```
power=1:0.01:5; f=10.^power;
```

(3) On the same graph, plot the magnitude of the experimentally measured transfer function in units of dB as a function of frequency. The experimental data should be plotted as discrete points and should include error bars that define the uncertainty in the data. The following command would perform this operation:

```
hold on; errorbar(fexperiment,20*log10(exptmag),erpos,erneg,'*'); hold off;
```

where `fexperiment` is an array containing the experimental frequencies used, `exptmag` is the magnitude of the transfer function and `erpos` and `erneg` are arrays that give the amplitude of the error bars in the positive and negative directions respectively (in units of dB). For example, if the uncertainty in the magnitude of the transfer function is `delta`, then the following commands would give the amplitude of the error bars in the positive direction and negative directions:

```
erpos=20*log10(exptmag+delta)-20*log10(exptmag);
erneg=20*log10(exptmag)-20*log10(exptmag-delta);
```

(4) Open a new figure window (you can use the command: `figure(2))` and plot the theoretical value of ϕ in units of radians as a function of frequency (equation 5). Again use a log scale for f. Calculate the uncertainty in the theoretical value of ϕ. If the uncertainties are not negligible add dotted lines to your graph showing the range of uncertainties, otherwise you can skip this step.

(5) On the same graph, plot phase of the experimentally measured transfer function. The experimental data should be plotted as discrete points and should include error bars that define the uncertainty.

Make sure to follow the instructions at the beginning of the manual to produce clear, informative, self-explanatory graphs. You can add a legend, title, and axis labels to your graphs using the following commands:

```
legend('Theory','Plus Uncertainty',' Minus Uncertainty', 'Experiment')
title('Low Pass Filter: Transfer Function Magnitude')
xlabel('Frequency [Hz]')
ylabel('Transfer Function Magnitude [dB]')
```

You can import the graph to a WORD documents by selecting Edit-> Copy Figure from the top menu bar and then simply pasting the figure (Ctrl v) directly into the document.

The expected cutoff (corner) frequency of the filter is defined as the frequency, where $v_{out} = \frac{1}{\sqrt{2}} v_{in}$, or $20\log(v_{out}/v_{in}) = -3$dB. From equation 4 it is easy to show that the cut off frequency is expected to be $f_c = \frac{1}{2\pi RC}$. Calculate the expected cutoff frequency for your circuit and by examining your graph verify that the filter is performing as expected.

Appendix 1:

Measurement of amplitude and phase

Fig. A1 represents an oscilloscope display and shows two voltage signals $v_1 = A_1 \cos(\omega t)$ and $v_2 = A_2 \cos(\omega t + \phi)$ at the same frequency. The angular frequency is $\omega = 10^5\ s^{-1}$, the frequency is $f = \omega/2\pi = 15.92$ kHz, the period is T = 1/f = 62.8 μ s and the amplitudes are $A_1 = 5\ V and\ A_2 = 7$V. The phase difference is $\phi = -1.2$ radians (or 68.8°). Note that when ϕ is negative, as here, then it corresponds to the sine wave being shifted to the right hand side on the display. Convince yourself that this is correct.[4]

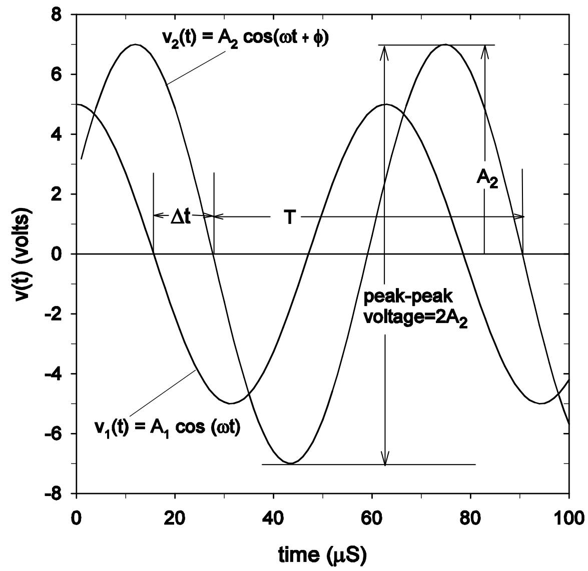


Fig. A1 Oscilloscope display

Amplitude

The amplitude A_2 is shown on the figure. To obtain the amplitude it is convenient to measure the peak-peak voltage, $2A_2$ in this case, as shown on the graph. This avoids problems if the display is not centred vertically.

Phase

The phase of $v_1(t)$ is ωt and that of $v_2(t)$ is $(\omega t + \phi)$. The difference in phase is ϕ and corresponds to the time difference Δt on the display. Note that when $\Delta t = T$ (the period) then $\phi = 2\pi$ radians (or 360°). Thus it follows that $\phi = 2\pi(\Delta t/T)$ radians or $360(\Delta t/T)$ degrees. On the graph $T = 62.8\,\mu$ s and $\Delta t = 12.0\,\mu$ s. Obviously, because Δt and T only appear as a ratio, they can be measured in any units as long as they are the same for both, e.g., seconds or centimeters on the display scale.

Note that it is very important that the 2 displays have the same zero on the vertical axis if Δt is to be accurately measured. To ensure this, switch each Channel input to GND and adjust the 2 horizontal traces to coincide with the scale zero line. Then switch both inputs back to AC.

[4] Note that the phase can be written $(\omega t - \phi)$. Both forms are commonly used.

Appendix 2: Breadboards

In the present context, a breadboard is a device which enables prototype circuits to be constructed and tested. The one you have is designed so that connections can be made to components without the need to solder them. An example is shown in Fig. A2. The components and wires are inserted into arrays of holes on the top surface of a plastic board. These holes are actually tiny sockets which are connected, as described below, underneath the board.

Notice that the columns of holes are in arrays which are either 2 or 5 columns wide. The 2 column arrays have all the holes in each *column* connected together. The 5 column arrays have all the holes on each *row* connected together. Test this by pushing wires into the holes and measuring the resistance between them with your ohmmeter. There are also 2 BNC connectors at the top RHS of the board which are used to bring signals to and from the board by coaxial cables. Each BNC connector has a wire socket at the side of it which can be used for making connections from the centre of the BNC socket to any point on the board. The outer part of the BNC socket is grounded and you can get a connection to this by using the screw-terminal ground connector on the board (the one next to the mains-power button).

Finally there are DC power supplies (+5 V, $\pm$15 V) built into the board.

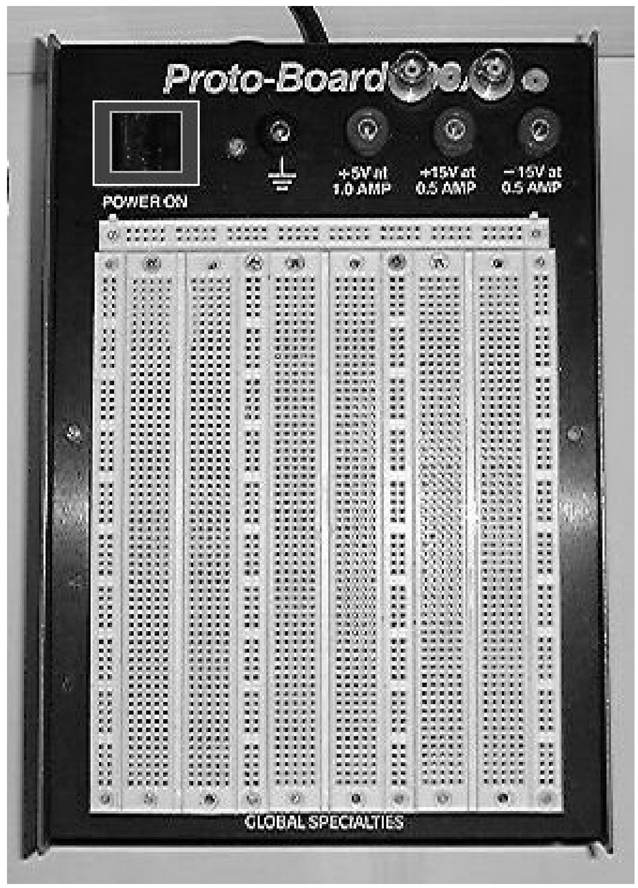


Fig. A2: A solderless breadboard

Appendix 3: An Easy Method for Calculating Uncertainties

The uncertainty Δy in the value of a function of several variables $y = f(a, b....)$ due to random measurement errors is given by the expression:

$$\Delta y = \left[\left(\frac{\partial f}{\partial a} \Delta a \right)^2 + \left(\frac{\partial f}{\partial a} \Delta b \right)^2 + ... \right]^{1/2} \qquad (1)$$

where Δa and Δb are the measurement errors associated with the variables *a* and *b*. Calculating the partial derivatives can sometimes be difficult and often leads to mistakes. If the measurements errors are relatively small then it is usually acceptable to use the following approximation for the derivatives:

$$\frac{\partial f}{\partial a} \approx \frac{f(a + \Delta a, b, ...) - f(a, b, ...)}{\Delta a} \qquad (2)$$

$$\frac{\partial f}{\partial b} \approx \frac{f(a, b + \Delta b, ...) - f(a, b, ...)}{\Delta b} \qquad (3)$$

Substituting these expressions into equation (1) gives a simplified formula for the uncertainty in the function:

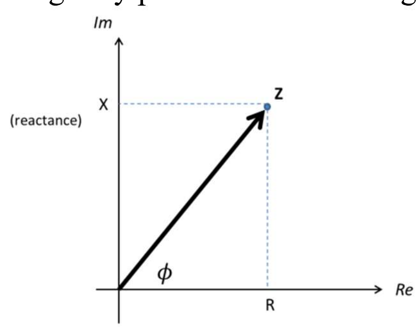
$$\Delta y \approx \left[\left(f(a + \Delta a, b, ...) - f(a, b, ...) \right)^2 + \left(f(a, b + \Delta b, ...) - f(a, b, ...) \right)^2 + ... \right]^{1/2} \qquad (3)$$

Electrical Impedance

The impedances of various circuit components are measured, both individually and in various combinations, using precision multimeters.

Introduction

When a sinusoidal voltage is applied to a linear circuit element, a sinusoidal current will flow through the element. The current will have exactly the same frequency as the sinusoidal voltage, but will not necessarily be in phase with the voltage. In order to analyze a linear AC circuit, you need to know how the amplitude of the current relates to the amplitude of the voltage and also how the phase of the voltage and current are related. An easy way to keep track of the amplitude and phase information is to define the voltage and current using complex numbers called 'phase vectors' or 'phasors'**.** If the phasors for the voltage and current are plotted as vectors in the complex plane, then the magnitude of the vectors represent the amplitude of the sinusoids and the angle between the vectors defines the phase angle between the sinusoids. The ratio of the phasors representing the voltage (**V**) and the current (**I**) are defined by a third complex number called the impedance **Z**=**V**/**I** where $\boldsymbol{Z} = R + jX$. We can visualize impedance on the real-imaginary plane as shown in Figure 1.



$$|Z| = \sqrt{R^2 + X^2}$$
$$\phi = tan^{-1}\left(\frac{X}{R}\right)$$

Figure 1: Impedance on real-imaginary plane

Electrical impedance extends the concept of resistance to AC circuits by taking the phase shift ϕinto account. This approach greatly simplifies the analysis of AC circuits but is only valid for linear circuits.

A linear circuit means that at a fixed input frequency, the ratio of potential difference to current is a constant independent of the amplitude of the voltage (or current). Also, no other frequencies besides the driving frequency appear in the current. Circuit components can be non-linear for various reasons. An inductor with a ferromagnetic core (e.g. iron or ferrite) may be non-linear due to effects of hysteresis, and the intrinsic non-linear relation between the magnetic flux and field exhibited by the core material. The experiment 'Ferromagnetic Hysteresis' deals with this in detail. Capacitors may be non-linear for similar reasons. Often the dielectric is a ferroelectric having analogous properties to a ferromagnet. Resistors are usually linear, but often diodes or other components are included specifically to make a circuit non-linear.

The present experiment investigates an inductor and a capacitor that were chosen to be linear. The circuits constructed from these will demonstrate the theory outlined in the next section.

Theory

When a time varying voltage such as $V\cos\omega t$ is applied to a circuit Ohm's Law is written as

$$\mathbf{V} = \mathbf{ZI} \qquad (1)$$

where **V, Z** and **I** are complex valued representations of the voltage, current and impedance. For most circuit elements **Z** varies with the frequency of the input and so too will the relationship between current and voltage.

In general, the current will be phase shifted relative to $\mathbf{V}$. This is accomplished using the appropriate vector representations for the impedance $\mathbf{Z}$ of a resistor, capacitor and inductor.

The current in a resistor is in phase with the voltage, meaning the resistance is simply a real number, i.e. a vector along the real axis, $\mathbf{Z_R}$ = R. Thus $\mathbf{V} = R\mathbf{I}$ and there is no phase shift between $\mathbf{V}$ and $\mathbf{I}$ as required.

An inductor with inductance L gives $V\cos\omega t = L(dI/dt)$ such that $I = (V/L)\int \cos\omega t dt = (V/\omega L)\sin\omega t = (V/\omega L)\cos(\omega t - \pi/2)$. Thus the current is phase shifted $-\pi/2$ relative to the voltage, or the current lags the voltage by $\pi/2$. Ideally magnitude and phase shift for the current is obtained if the impedance of the inductor is taken to be $j\omega L$, i.e. purely imaginary (Figure 2). Real inductors, however, have resistance R_L (Figure 3) so in general the impedance of an inductor is taken to be

$$\mathbf{Z_L} = R_L + j\omega L \qquad (2)$$

In a similar manner it is found that a capacitor with capacitance C has impedance

$$\mathbf{Z_C} = 1/j\omega C \qquad (3)$$

or equivalently $-j/\omega C$ (Figure 4). In this case the current leads the voltage by π/2.

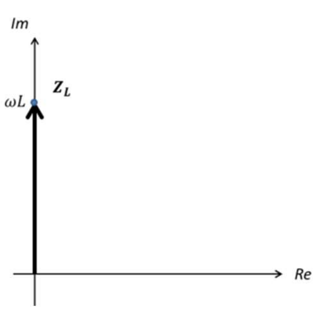


Fig 2: Impedance of ideal inductor

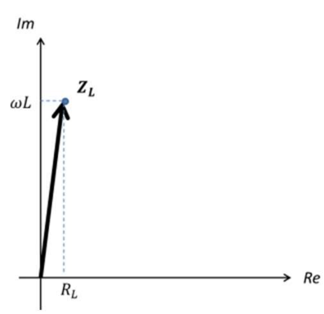


Fig 3: Impedance of real inductor

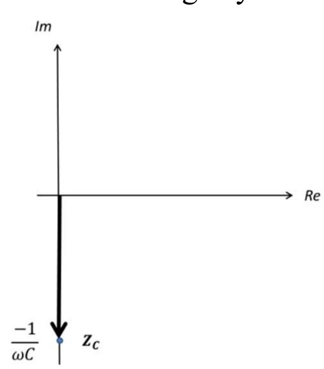


Fig 4: Impedance of capacitor

It can be shown that any circuit can then be analyzed using the usual circuit laws (Kirchhoff, Thévenin, etc.) provided impedances are added as vectors[5]. For circuit elements with impedance $\mathbf{Z} = R + jX$ the magnitude of the impedance is

$$|\mathbf{Z}| = Z = \sqrt{R^2 + X^2} \qquad (4)$$

From equation (1) $|\mathbf{Z}|=|\mathbf{V}|/|\mathbf{I}|$ but $|\mathbf{V}|/|\mathbf{I}|$ = Vrms/Irms so experimentally the magnitude of the impedance can be found using

$$|\mathbf{Z}|=\text{Vrms/Irms} \qquad (5)$$

where Vrms and Irms are the RMS voltage and current across the inductor respectively.
Since the reactive part ($X=\omega L$) of the inductor vanishes at DC ($\omega=0$), the resistance of the inductor can be measured using an ohmmeter. Once the resistance and magnitude of the impedance are known, X can be determined from equation (4).

Experiment

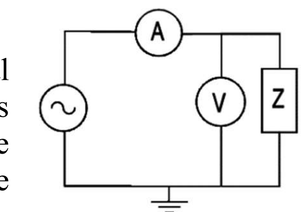


Fig. 1 Measuring RMS current and voltage to determine impedance

The circuit for measuring impedance is shown in Figure 1. The digital multimeters (DMM) should be set to AC. On the AC setting the instruments measure RMS values. The common terminal of the voltmeter should be connected to ground. Note that all phase information is lost in the measurement.

The accuracy of meters is typically specified as $\pm$ a percentage of the input $\pm$ a number of digits in the last place. The specifications are usually on a sheet pasted to the underside of the meters, but if it is not available assume an accuracy of $\pm 0.5\% \pm 2$ digits for AC volts, $\pm 0.75\% \pm 2$ digits for AC current, and $\pm 0.2\% \pm 2$ digits for (DC) resistance. (The accuracy for DC voltage and current will be much better but these scales are not used). Because there will be slight differences between the calibration factors on different scales, it is usually better to stay on one scale for a series of measurements.
Example: Let's say you measured 3.04 V. The last digit shown on the meter is in the hundredths so "2 digits" would be 0.02 V. The uncertainty would be +/- (0.005*3.04 + 0.02)V.

A powerful technique for reducing the uncertainty in a measurement is to set up the experiment so that the required value is either the slope or intercept of a regression line. In part 1 of this experiment we wish to determine L and R_L for an inductor and C for a capacitor. For an inductor, we know that $X= \omega L$ and so from equation (4) we get $|\mathbf{Z}|^2=R^2+ \omega^2 L^2$. Equation (5) indicates that we can experimentally determine $|\mathbf{Z}|^2$ as $|Z|^2 = (\text{Vrms/Irms})^2$. If $|\mathbf{Z}|^2$ is plotted as a function of ω^2, the slope of the regression line will give the value of L^2 and the intercept will give R^2. Similarly, for a capacitor, $|\mathbf{Z}|^2=R^2+ 1/(\omega^2 C^2)$. Here, if $|\mathbf{Z}|^2$ is plotted as a function of $1/\omega^2$, the slope will give $1/C^2$ and the intercept R^2 which for an ideal capacitor should be zero within uncertainty. For this part of the experiment measurements of the RMS voltage (Vrms) and current (Irms) across the element over a range of frequencies but at a constant source voltage are required.

[5] The complex math can be dealt with easily in Matlab or on your calculator. Instructions for doing complex math with the CASIO calculator can be found in the reference section of the course webpage.

In part 2 a good estimate of the magnitude of the impedance of parallel and series combinations of the inductor and capacitor at a fixed frequency will be made. From equation (5) it can be seen that if you plot Vrms as a function of Irms, the slope of the regression line will give the magnitude of the impedance $|\mathbf{Z}|$ at that frequency. In this case, keeping the frequency constant, the RMS voltage and current will be measured over a range of source voltages.

Part 1: Using a fixed source voltage while varying frequency to determine L, R_L and C

a) In this part you will measure RMS current and voltage as a function of frequency in order to determine the magnitude of the impedance as a function of frequency. Start by measuring the RMS voltage and current for the inductor. Set the signal generator to a frequency of approximately 200 Hz and adjust the output voltage so that it is as large as possible while keeping $I_{RMS} < 100$ mA which is the maximum current rating for the inductor. Measure the RMS current and voltage for the following frequencies: f=[200 490 663 800 916 1020 1114 1200] Hz. Note: you don't need to use these frequencies exactly, however, the frequencies shown should produce an approximately equal spacing between the data points when $|\mathbf{Z}|^2$ plotted as a function of ω^2. Make sure you record the exact value of the frequencies you use.

b) Repeat this experiment using the capacitor. Use the following frequencies: f=[208 220 235 255 280 310 350 420 560 1200]. Again, you don't need to use these frequencies exactly but the frequencies given should produce approximately equally spaced points when you plot $|Z|^2$ and a function of ω^{-2} .

c) Use the DMM to measure R_L and C so you have an idea of what to expect when you analyse the data.

Part 2: Using a varying source voltage to determine Z at a fixed frequency

Choose a frequency near the upper range of frequencies used for part 1 and record the frequency selected. Collect Vrms and Irms data using a range of source voltages (up to the maximum voltage found in part 1) for a) the parallel combination of the inductor and capacitor, and b) the series combination of the inductor and capacitor. Vary the voltage from zero up to the maximum possible. Take a total of about 10 voltage and current readings for each impedance under test. This should give good statistics in the analysis. For case a) (parallel combination) once you reach the maximum voltage, measure the current through the inductor alone and also through the capacitor alone. Similarly, for case b) (series combination) at the maximum voltage measure the voltage across the capacitor and the voltage across the inductor.

Analysis and Report

A. Using the data from part 1a), plot $|Z|^2$ as a function of ω^2. Use the slope and intercept to calculate L and R_L from equation (4).
As a general rule, in any analysis it to recommended to take into account the accuracy of each data point, but typically all points are given equal weighting in a computational fit. The present data provide an example where the uncertainties vary (even if the uncertainty on Z is constant the uncertainty on $|Z|^2$ won't be) so you should use a weighted fit (for more information see section on linear regression in the error analysis section).

Using the data from 1b), plot $|Z|^2$ as a function of ω^{-2} and use the slope to calculate C. If you have used a weighted regression you will likely find that the intercept is 0 to within the uncertainty.

How do C and R_L compare to the values measured using the DMM?

B. Graph the data for parts 2 a) and b). Determine the magnitudes of the impedances Z at your chosen frequency as well as the uncertainty using least square fits to the data. In this case since the impedance isn't squared the variation in uncertainty is relatively small and so an unweighted fit is appropriate (see pp16-17).

C. Calculate the expected impedances for parts 2 a) and b) at the test frequency chosen using the values of L, R_L and C, determined in part A (write each impedance in the form $\mathbf{Z} = R + jX$). Calculate the magnitude of each impedance. How do the magnitudes of the impedances compare to the experimental results from B?

The following MATLAB commands would give the impedances and the magnitudes of the impedances for the parallel combination `(Zp and magZp)` and series combination `(Zs and magZs)`

```
ZL=RL+j*w*L;
ZC=1/(j*w*C);

Zp=1/(1/ZL+1/ZC)
magZp=abs(Zp)

Zs=ZL+ZC
magZs=abs(Zs)
```

D. Using the impedances calculated in C and the measured voltage, calculate the expected currents I_L and I_C for last data point in the parallel configuration (part 2a) using equation (5). Compare with experimentally measured currents.

E. Using the impedances calculated in C, calculate the expected voltages V_L and V_C for the last data point in part 2b). (This can be done using the measured current in equation (5) or by recalling that impedances in series act like a voltage divider.) Compare with the experimentally measured voltages.

h/e and the Photoelectric Effect

Introduction

The emission and absorption of light was an early subject for investigation by the German physicist Max Planck. As Planck attempted to formulate a theory to explain the spectral distribution of emitted light based on a classical wave model, he ran into considerable difficulty. Classical theory (Rayleigh-Jeans Law) predicted that the amount of light emitted from a black body would increase dramatically as the wavelength decreased, whereas experiment showed that it approached zero. This discrepancy became known as the ultraviolet catastrophe.

Experimental data for the radiation of light by a hot, glowing body showed that the maximum intensity of emitted light also departed dramatically from the classically predicted values (Wien's Law). In order to reconcile theory with laboratory results Planck was forced to develop a new model for light called the quantum model. In this model light is emitted in small discrete bundles, or *quanta*.

Associated with such quanta is a new fundamental constant of nature, *h*, known as Planck's constant. In this experiment you will accurately determine the ratio of h to the magnitude of the charge of an electron *e*. This will establish *h* in units of eV·s. The experimental technique used is known as the photoelectric effect.

Theory

Planck's Quantum Theory

In 1901 Planck published his law of radiation. In it he stated that an oscillator, or any similar physical system, has a discrete set of possible energy values, or levels; energies between these values never occur.

Planck went on to state that the emission and absorption of radiation is associated with transitions, or jumps, between two energy levels. The energy lost or gained by the oscillator is emitted or absorbed as a quantum of radiant energy, the magnitude of which is expressed by the equation:

$$E = h\nu \quad (1)$$

where E equals the radiant energy, ν is the frequency of the radiation, and h is a fundamental constant of nature. The constant, h, became known as Planck's constant.

Planck's constant was found to have significance beyond relating the frequency and energy of light, and became a cornerstone of the quantum mechanical view of the subatomic world. In 1918 Planck was awarded a Nobel prize for introducing the quantum theory of light.

The Photoelectric Effect

In photoelectric emission, light strikes a material causing electrons to be emitted. The classical wave model predicted that as the intensity of the incident light was increased the amplitude, and thus the energy, of the wave would increase. This would then cause more energetic photoelectrons to be emitted. The new quantum model, however, predicted that higher frequency light would produce higher energy photoelectrons, independent of intensity, while increased intensity would only increase the number of electrons emitted (the so-called photoelectric current). In the early 1900s several investigators found that the kinetic energy of the photoelectrons was dependent upon the wavelength, or frequency, and independent of intensity, while the magnitude of the photoelectric current, or number of electrons, was dependent upon the intensity as predicted by the quantum model. Einstein applied Planck's theory and explained the photoelectric effect in terms of the quantum model using his famous equation:

$$E = h\nu = \mathrm{KE}_{\max} + W_0 \quad (2)$$

where KE_{max} is the maximum kinetic energy of the emitted photoelectrons, and W_0 is the energy needed to remove them from the surface of the material (the work function). E is the energy supplied by the quantum of light known as a *photon*. Based largely on the success of this relation, Einstein received the Nobel prize in 1921.

The h/e Experiment

A photon with energy $h\nu$ is incident upon an electron in the cathode of a vacuum tube. The electron uses a minimum W_0 of its energy to escape the cathode, leaving it with a maximum energy of KE_{max} in the form of kinetic energy. Normally the emitted electrons reach the anode of the tube, and can be measured as a photoelectric current. However, by applying a reverse potential V between the anode and the cathode, the photoelectric current can be stopped. KE_{max} can be determined by measuring the minimum reverse potential needed to stop the photoelectrons and reduce the photoelectric current to zero.* Relating kinetic energy to stopping potential gives the equation:

$$\mathrm{KE}_{\max} = Ve \quad (3)$$

Therefore, using Einstein's equation,

$$h\nu = Ve + W_0 \quad (4)$$

When solved for V, the equation becomes

$$V = (h/e)\nu - (W_0/e) \quad (5)$$

If we plot V versus ν for different frequencies of light the V intercept is equal to $-W_0/e$ and the slope is h/e (this is the value of h in eV·s. This directly establishes the numerical value for Planck's constant. Using the accepted value for e, 1.602×10^{-19} C,we can determine Planck's constant, h in J ·s

---

* Note: In experiments with the PASCO h/e apparatus, the stopping potential is measured directly rather than by monitoring the photoelectric current. See the Appendix at the end of these instruction sheets for details.

Experiment

In experiments with the *h*/*e* apparatus, monochromatic light falls on the cathode plate of a vacuum photodiode tube that has a low work function, W_0. Photoelectrons ejected from the cathode collect on the anode.

The photodiode tube and its associated electronics have a small capacitance which becomes charged by the photoelectric current. When the potential on this capacitance reaches the stopping potential of the photoelectrons the current decreases to zero, and the anode-to-cathode voltage stabilizes. This final voltage between the anode and the cathode is therefore the stopping potential of the photoelectrons.

To let you measure the stopping potential, the anode is connected to a built-in amplifier with and ultra-high input impedance ($\approx 10^{12}\ \Omega$), and the output from this amplifier is connected to the output jacks on the front panel of the apparatus. This high impedance, unity gain ($V_{out}/V_{in} = 1$) amplifier lets you measure the stopping potential with a digital voltmeter.
Note: While the impedance of the unity gain amplifier is very high, it is not infinite and some charge leaks off. Charging the apparatus is analogous to filling a bath tub while the drain is partly open. The true stopping potential may not be reached if it is charged slowly compared to the rate it leaks (the tub never gets full).

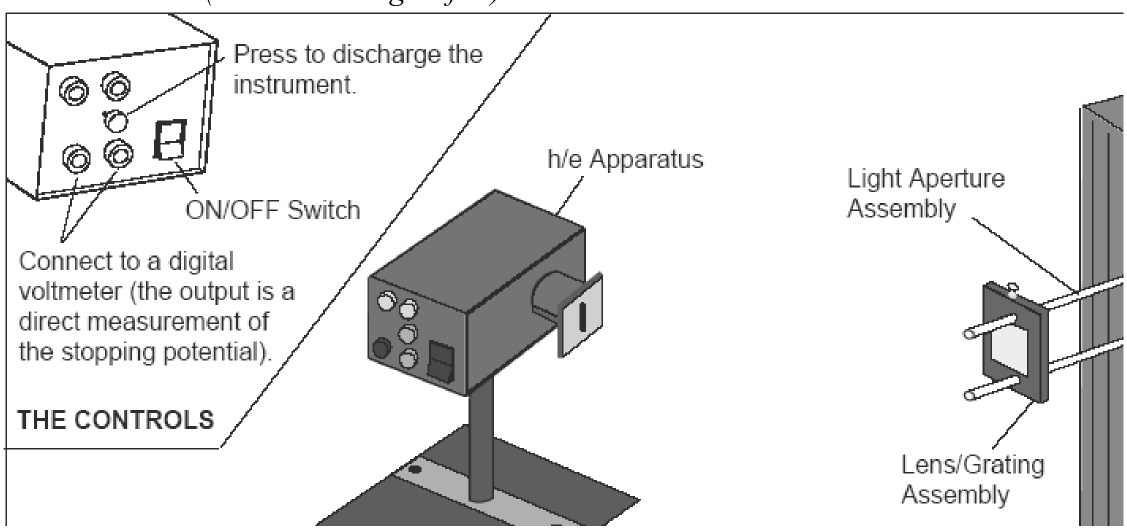


Figure 1: h/e apparatus with mercury vapour light source (DVM not shown)

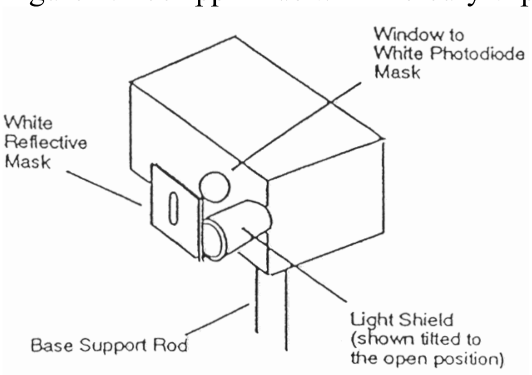


Figure 2: h/e apparatus with light shield tilted

A: Setup

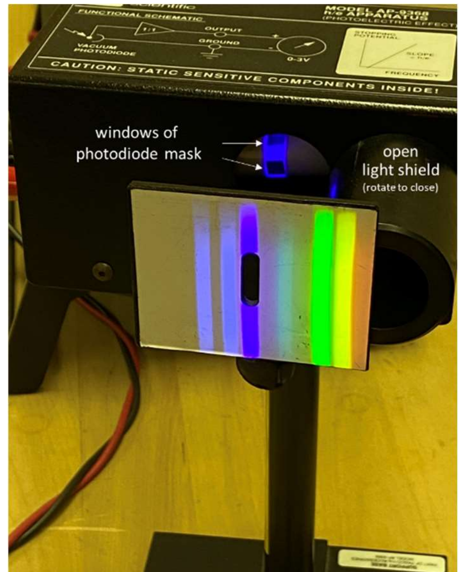


Fig 3. Align light with windows of photodiode mask

Turn on the Hg vapour light source, dim the lights, and wait for several minutes. Focus the light emitted from the Hg source onto the slot in the white reflective mask on the *h*/*e* apparatus. Note that the lens/grating may be flipped over with a considerable loss in intensity – use the arrangement which yields the highest intensity. Also check the spectral lines to the left and right of center and use the ones that are clearer. Tilt the Light Shield of the apparatus out of the way to reveal the white photodiode mask inside the apparatus. Slide the lens/grating assembly forward and backward on its support rods until you achieve the sharpest image of the aperture centered on the windows in the photodiode mask. Secure the lens/grating by tightening the thumbscrew.

Align the system by rotating the *h*/*e* apparatus on its support base so that the same colour light that falls on the opening of the light screen falls on the windows in the photodiode mask, with no overlap of colour from other spectral lines as shown in Figure 3. Return the Light Shield to its closed position.

Check the polarity of the leads from your digital voltmeter (DVM), and connect them to the OUTPUT terminals of the same polarity on the *h*/*e* apparatus.

Use a piece of paper as a screen to look at the spectral lines on both sides (left and right) and decide whether those to the left or the right are more distinct and use the lines on the side that are most distinct for the rest of the experiment.

B: Effect of light intensity

1. Adjust the *h*/*e* apparatus so that only one of the spectral colours falls upon the opening of the mask of the photodiode. If you select the green or yellow spectral line, place the corresponding coloured filter over the White Reflective Mask on the *h*/*e* apparatus.

2. Place the Variable Transmission Filter in front of the White Reflective Mask (and over the coloured filter, if one is used) so that the light passes through the section marked 100% and reaches the photodiode. Record the stopping potential from the DVM in a table. Press the instrument discharge button, release it, and record how much time is required to recharge the instrument to the maximum voltage.

3. Move the Variable Transmission Filter so that the next section of it is directly in front of the incoming light. Record the % transmission, stopping potential, and the time to recharge after the discharge button has been pressed and released.

4. Repeat Step 3 until you have tested all five sections of the filter. Is the expected trend observed?

5. Repeat the procedure using a second colour from the spectrum.

C: Energy of photons as a function of frequency

1. You can identify five colours in the mercury light spectrum. Adjust the *h*/*e* apparatus so that only the yellow coloured band falls upon the opening of the mask of the photodiode. Place the yellow coloured filter over the White Reflective Mask on the *h*/*e* apparatus.

2. Record the DVM voltage reading (stopping potential) in the table below (zero the *h*/*e* apparatus before each measurement, and allow for the full accumulation of charge before you record your measurement).

3. Repeat the process for each colour in the spectrum. Be sure to use the green (yellow) filter when measuring the green (yellow) spectrum.

Light Colour	**Frequency (Hz)**	**Stopping Potential**
Yellow	5.19×10^{14}	
Green	5.49×10^{14}	
Blue	6.88×10^{14}	
Violet 1	7.41×10^{14}	
Violet 2	8.20×10^{14}	

Analysis

Make a graph of the time to reach the stopping potential as a function of transmission. Describe the effect that passing different amounts of the same coloured light through the Variable Transmission Filter has on the charging time after pressing the discharge button. Come up with a model for the relationship between transmission and time. Does your data support this model? (You can test your model by trying to create a linear graph based on the model.)

Do your results support the quantum model prediction that increased intensity would increase the number of photoelectrons emitted (the so-called photoelectric current) but not the energy of the photoelectrons?

Make a graph of stopping potential V vs. frequency ν. Describe the effect that different colours of light had on the stopping potential, and thus the maximum energy of the photoelectrons. From the linear regression determine *h*/*e*, and thus *h*, (see eqn. 5). Compare your experimental value with the accepted value.

Do the results support the quantum model prediction that higher frequency light would produce higher energy photoelectrons.

Bending Beams and Strain Gauges

The strains in a cantilever beam subject to loading are examined by means of strain gauges. Young's modulus and Poisson's ratio for the beam material are determined.

Introduction and Theory

(i) Bending Beams

When a beam of rectangular cross section is bent to a radius of curvature R, the convex surface is stretched along the longitudinal axis while the concave surface is compressed. For reasonably thin beams, the longitudinal strains ε_ℓ at corresponding points on these two surfaces are equal in magnitude, but opposite in sign. For a beam of thickness a, the strain is given by

$$\varepsilon_\ell = \frac{a}{2R} \quad (1)$$

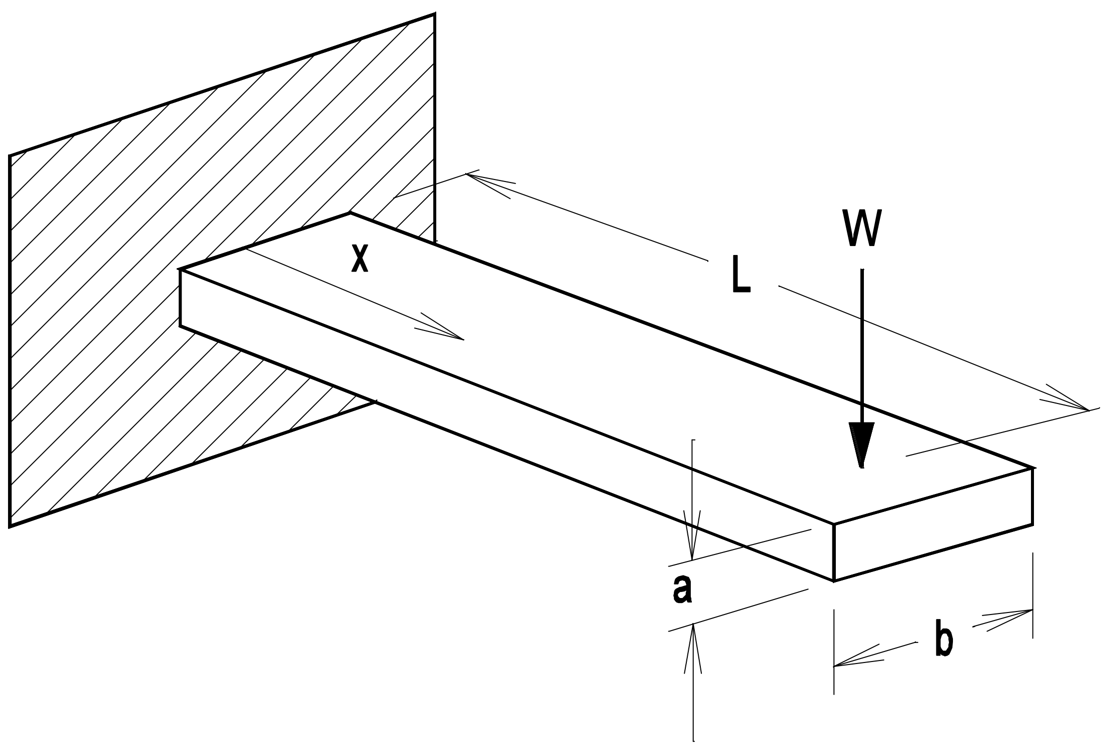


Figure 1. Loaded cantilever beam

Fig. 1 shows a cantilever beam of thickness a, and width b, clamped at one end and depressed by a force W, a distance L from the clamp. The torque on the beam exerted by the load is $W(L-x)$, at a distance x from the clamped end. The internal stresses which accompany the strains in the beam provide a balancing torque, giving the equation

$$W(L-x)=\frac{YI_A}{R(x)} \quad (2)$$

where I_A is called the second moment of area. I_A is a geometrical property of the beam that measures the beams ability to resist bending. In this particular case, $I_A = a^3b/12$. In the present experiment it is often convenient to write the results in terms of the depression of the bar at the point where W is applied, say y. Calculation shows

$$y=\frac{WL^3}{3YI_A} \quad (3)$$

where Y is Young's modulus of the material. Using these equations one finds

$$\varepsilon_\ell = \frac{3a(L-x)y}{2L^3} \quad (4)$$

(ii) Transverse Effects

When a beam is subject to uniaxial stress (i.e. stress applied in one direction only), not only is the beam deformed (strained) along the longitudinal direction, it also exhibits deformation (strain) of opposite sign in both perpendicular directions. The negative of the ratio of the transverse strain to the longitudinal strain is called Poisson's ratio σ, i.e.

$$\sigma = -\frac{\varepsilon_t}{\varepsilon_\ell} \quad (5)$$

This effect is present with the loaded cantilever beam. The top surface has a negative ε_t (bends up at both edges), and the bottom surface has a positive ε_t (also bends up at both edges). One can relate the transverse strains of the top and bottom surfaces to the local ε_ℓ by Eq. 5. The sides of the beam also exhibit deformations due to strain, but these effects are not examined here.

(iii) Strain Gauges

Strain is conveniently measured with a strain gauge consisting of a thin layer of metal or semiconductor deposited on a thin insulating film, typically mylar. The underneath of the insulator is bonded directly to the sample such that the gauge and the beam bend equally. Bending causes the resistance of the film to change as a function of the applied strain. Commercial gauges are supplied with a calibration factor, called the gauge factor, defined by $GF = (\delta R/R)/\varepsilon_\ell$, where $\delta R/R$ is the relative change in the resistance caused by the strain ε_ℓ. The gauge factor of a commercial gauge is typically around 2.0.

An examination of any of the gauges in this experiment shows that the metal film is in the form of a wire grid with the wires running mainly in one direction, say the major axis. The gauge is designed to be highly sensitive when stretched in the direction parallel to the major axis, the direction in which the majority of the wires lie. The ratio of the two sensitivities is called the transverse sensitivity K_t, a dimensionless quantity, typically about 0.01.

During calibration of a gauge, it is bonded to a sample that is strained uniaxially. This produces both longitudinal and transverse strains, both of which affect the gauge resistance, though longitudinal strain has a much greater effect. If the gauge is subsequently used to measure strain in a material with a different Poisson's ratio, GF will be slightly different. Because K_t is very small, this is a tiny effect and can usually be ignored.

There is, however, a related effect. When the transverse strain ε_t is to be measured, the gauge must be oriented with the major axis parallel to ε_t. Under these conditions, the gauge is subject to a much larger strain along the minor axis, causing the GF to appear significantly in error. If ε_ℓ and ε_t' represent the measured longitudinal and the apparent transverse strains, then the true transverse strain is given by

$$\varepsilon_t = \varepsilon_t' - K_t \varepsilon_\ell \qquad (6)$$

This is only a first order correction, but it is accurate to better than 1% and is sufficient for the purposes if this lab. However, it is imperative that the relative sign of the correction is accurate in the application of Eq. 6.

Experiment

The apparatus includes a set of strain gauges mounted on two standard aluminum bars, a beam deflection apparatus, and a highly sensitive Wheatstone bridge to measure the changes in gauge resistance. Figure 2a) shows the first beam that you will use that has 3 strain gauges. The two red wires are both attached to the right side of all the gauges. The white, black and green wires are each attached to the left side of one of the strain gauges. The deflection apparatus and Wheatstone bridge are shown in figure 2b)

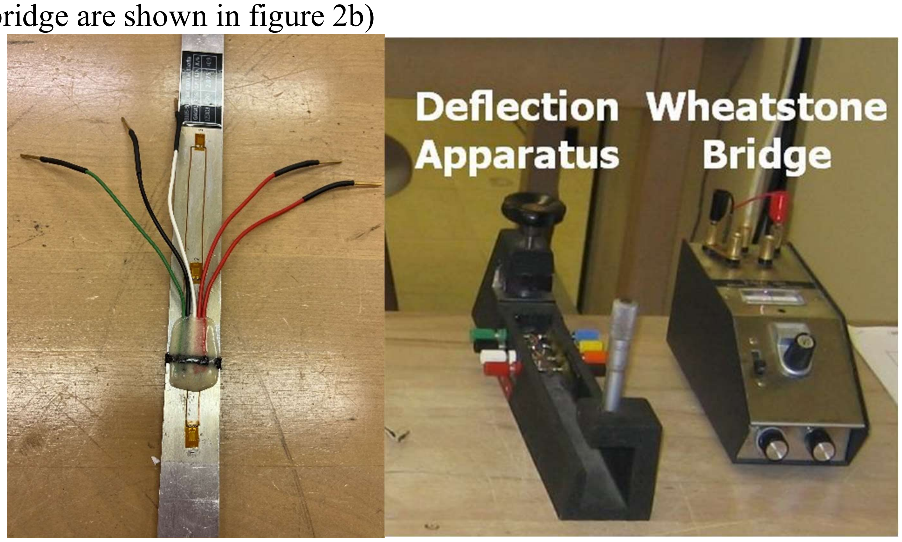


Figure 2: a) Beam with 3 strain gauges b) Deflection apparatus and Wheatstone bridge

The Wheatstone bridge is described in the information package provided in the lab.

The beams can be stressed either by micrometer depression or by loading with weights. Because the strain is a linear function with respect to distance x (see Eq. 4) the gauges may be taken to be located at the center of the wire grids. There are small v's marked on the sides of the gauge indicating the location of the centre. You should measure the beam dimensions, distances of the gauges to the clamp and distance of the two possible load points to the clamp[6].

The Wheatstone bridge is contained in a module with an external power supply. This must be turned OFF whenever circuits are being connected or disconnected. The left hand switch on the bridge module must be permanently set to OFF.

When a beam is clamped into the deflection apparatus, connect the strain gauge wires to the binding posts on the side (refer to Fig. 3 for wiring diagrams). Be sure all connections are secure, as the resistance changes are very small in this experiment. Note GF and K_t for the gauge and set the left dial on the lower edge of the bridge module to the value of GF.

Plug the red power lead into the P+ post, and the black power lead into the P- post on the module, leaving the power switch OFF. Connect the strain gauge wires to the module. Note that the connections differ for the measurement of tensile and compressive strain.

When all connections have been made and checked, turn the power supply ON. Now zero the bridge. Adjust the zero (the bottom right hand knob on the module, labelled BALANCE) so that the needle is in the null position when the strain dial (the large central knob with the digital counter) is at 0. After a load is applied adjust the strain dial until the needle is back in the null position. The counter will now read ε directly in units of $10^{-5} = 10\,\mathrm{m}\varepsilon$. The dimensionless 'unit' $\mathrm{m}\varepsilon$ is called microstrain. (The maximum reading of the dial is 1000 units = 10 000mε).

[6] An explanation of how to read the Vernier calipers can be found in the Ultrasound lab or in the video https://stream.queensu.ca/Watch/t3N8HgSf

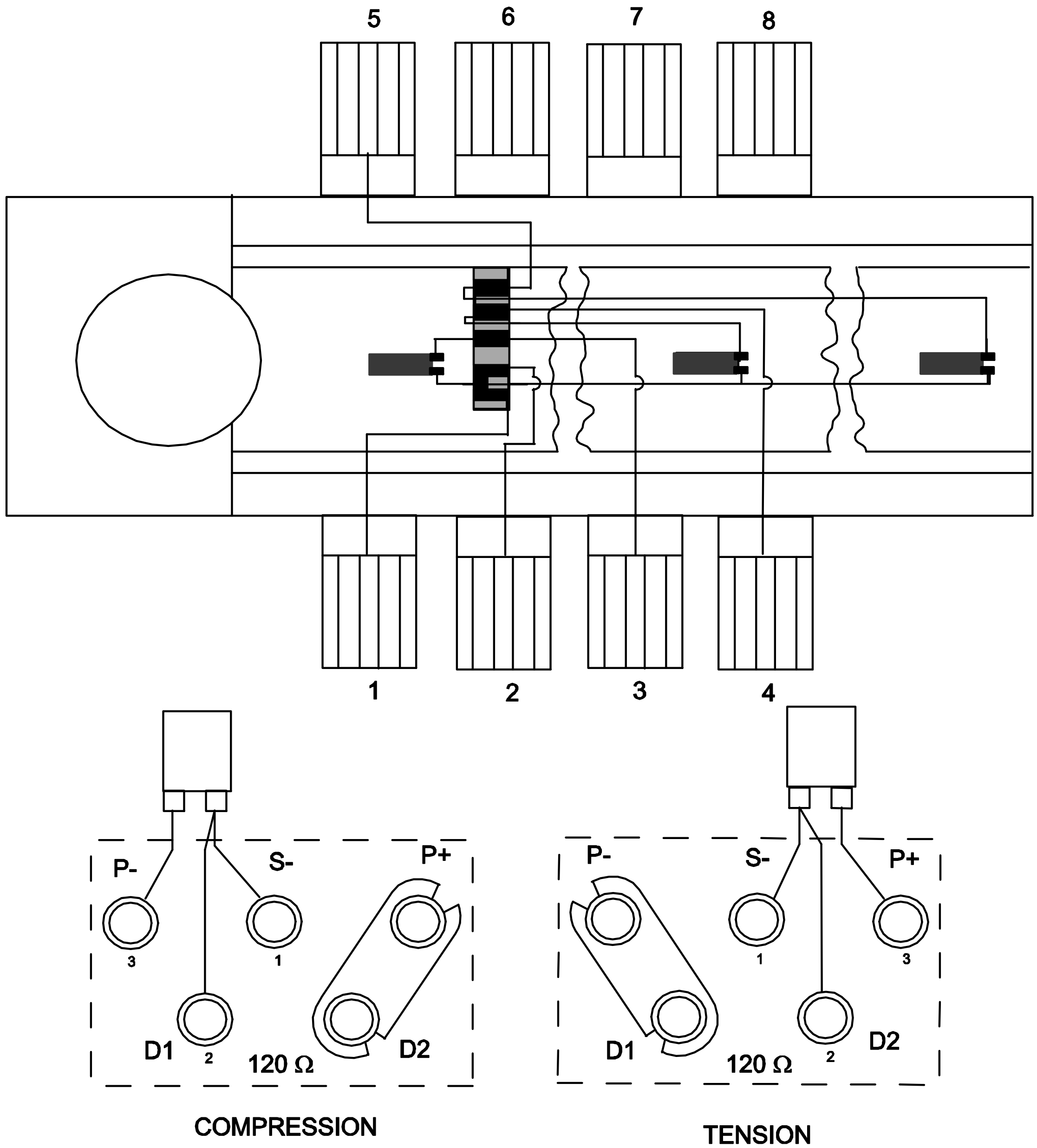


Figure 3. Wiring Diagram. The two red wires go to S- and one of the other wires goes to P+ or -

1. Strain measurements using micrometer deflection

Insert the triple-gauge beam into the deflection apparatus with the strain gauges on the upper surface. Connect the strain gauge closest to the clamp and measure ε_ℓ for five equal increments of depression y up to a maximum of 0.5". Estimate the errors in these measurements. Plot ε_ℓ versus y as data is being collected. The plot should be linear according to Eq. 4.

Switch off the power supply. Connect the next gauge along the beam by connecting the correct strain gauge wire to the module (see information package provided). Re-zero the bridge. Repeat the strain measurements for the same five micrometer positions, and plot the results on the same graph. Finally, repeat procedure for the third strain gauge. Measure the distance x from the clamp to the center of the strain gauges, and also to the micrometer L. Measure the thickness a, and the width b, of the beam.

2. Strain measurements using loading to determine Young's modulus

Reconnect the strain gauge nearest the clamp and re-zero. Weigh the weight-hanger and place it in the dimple at the end of the beam. Measure ε_ℓ for five weights (W) up to a maximum load of about 5 kg. Make a quick plot of ε_ℓ as a function of W. According to Eqs. 3 and 4 this should again be a linear relationship. Combine equations 3 and 4 to eliminate y and obtain an expression for Y as a function of *W*. Obtain an estimate of Y from the slope of the graph and compare your result with the accepted value Young's modulus of aluminium. If you aren't within 10% check your measurements. Measure L which, in this case, is the distance from the clamp to the dimple.

3. Comparison of strain on opposite sides of beam

It is expected that for thin beams the strain on the top and bottom will be equal but opposite in sign i.e. one side will be in extension and the other in tension. In order to investigate this you will turn over the beam and measure the strain on the bottom of the beam at the first gauge and compare it to the measurements made when that gauge was on the top. Switch off the power supply, and invert the bar so that the strain gauges are on the bottom surface. Using the strain gauge nearest the clamp, confirm that resistance change in now in the opposite sense to those in the previous measurements, i.e. the resistance now decreases with increasing depression of the bar, meaning strain is compressive rather than extensive. Reconnect the circuit to measure compressive strain. Repeat the set of five measurements of ε_ℓ as a function of y up to 0.5".

4. Poisson's ratio

Insert the beam with a strain gauge on both the top and bottom surface. Note that the orientations are perpendicular to one another so that one will measure ε_ℓ and the other ε_t. Measure the distances x from the clamp to the two gauges. If these are not identical you will need to make an allowance in the analysis. Also measure a and b for the beam.

Recalibrate and re-zero the Wheatstone bridge and make five measurements of strain as a function of y for the top gauge. Repeat for the bottom gauge. As mentioned earlier, when a beam is subject to uniaxial stress the beam strained along the longitudinal direction and exhibits strain of opposite sign in the perpendicular direction. Explain why you observe both strains to be in the same direction.

Analysis and Report

1. Plot ε_ℓ versus y for each of the three gauges. Perform linear regression and compare the slopes with the predicted slopes from Eq. 4.

2. Plot ε_ℓ as a function of *W*. Perform linear regression and determine Y from the slope using the expression for Y as a function of *W* that you obtained from equations 3&4. Compare your result with the accepted range for Young's modulus of aluminium. Remember to reference your accepted values.

3. Plot ε_ℓ versus y for the compression data from part 3 (gauge on the bottom of the beam). Perform linear regression and compare the slope to that for the extension data from part 1 (gauge on the top of the beam). Do they agree?

4. For the beam with longitudinal and transverse strain gauges correct the transverse data using Eq. 6. Be careful of the signs in this equation as it assumes ε_ℓ and ε_t are measured at the same point. on the same side of the beam. Plot both strains as a function of *y*. Determine the Poisson's ratio (σ) for the beam from the ratio of the slopes (see equation 5) and compare with the accepted value for aluminium. Remember to reference your accepted value.

Young's modulus of steel

In this experiment, Young's modulus of steel wire is measured by two different methods.

Introduction and Theory

Hooke's law states that the ratio of stress to strain is a constant for small deformations. The constant is a coefficient of elasticity of the material. There is a different coefficient of elasticity for each different kind of strain. When the strain is an extension in one direction, this coefficient is called Young's modulus. Consult your first year text for an introduction to elasticity and Young's modulus.

The stress in the present case is provided by a tensile force F acting over the cross-sectional area A. The resulting change in length $\Delta\ell$ is equivalent to a strain $\Delta\ell/\ell$, where ℓ is the unstretched length. Hooke's law for this situation reads

$$E = \frac{stress}{strain} = \frac{F/A}{\Delta\ell/\ell}. \quad (1)$$

where E is Young's modulus.

The law is obeyed by most materials when the strain is not too large. Most materials can experience strains of 0.1-1% before entering the plastic regime in which they will be permanently deformed and will no longer obey Hooke's law. At higher stress levels the relationship between the stress and the strain becomes non-linear and the concept of Young's modulus breaks down. In this region the material becomes permanently deformed by the stress with irreversible internal damage. The change over occurs at a level of stress called the *yield stress*, the value of which is very dependent on the material.

In this experiment Young's modulus will be measured by two different methods. The first method we will use is called a *static* method because the stress is applied to a wire and the system is allowed to come to equilibrium. The second method is called a *dynamic* method because the wire under investigation is made part of an oscillatory system.

The dynamic method is based on the following situation. A wire of length ℓ and cross-sectional area A is hung vertically with a mass M attached to its lower end. If the mass is displaced a vertical distance x, the strain is x/ℓ. From equation (1) we see that the restoring force on the mass M is $-(EA/\ell)x$. This gives the mass an acceleration

$$\ddot{x} = -(EA/M\ell)x. \quad (2)$$

This is the equation of simple harmonic motion. Hence the mass will vibrate simple harmonically, i.e.,

$$x = x_o \cos(\omega t + \phi) \quad (3)$$

Taking the second derivative gives the acceleration as

$$\ddot{x} = -\omega^2 x_o \cos(\omega t + \phi) \quad (4)$$

Substituting x and $\ddot{x}$ from equations (3) and (4) into (2) gives a frequency f given by

$$2\pi f = \omega = \sqrt{\frac{EA}{M\ell}} \qquad (5)$$

Such oscillatory effects are regularly observed when elevators have very long cables as in tall buildings or mine shafts. Note that transverse vibrations can also be set up with the same arrangement but their frequency is essentially independent of E (see *Mechanical Resonance*). Also note that the above derivation assumes that M is much larger than the mass of the wire. In principle the massive weight is not required at all; the wire alone will vibrate longitudinally, but this happens at much higher frequencies.

Experiments

The following figure shows the arrangement of both experiments.

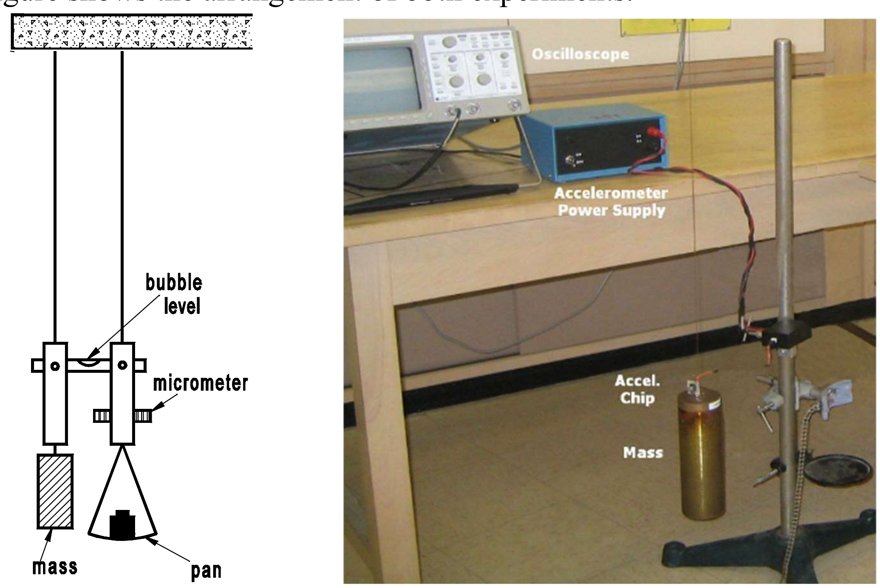


Fig.1 Static (left) and dynamic (right) apparatus

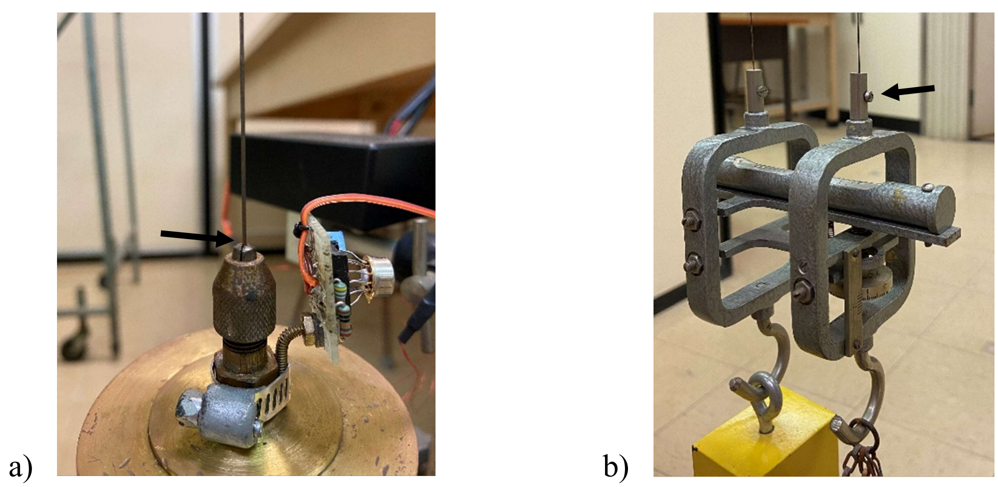


Fig. 2 Measuring length of wire: wire connections a) connection at the top of both wires and bottom of dynamic setup b) connection at bottom of static setup is with the setscrew seen here.

(i) Static method
Two vertical wires are fastened to a rigid support on the ceiling. One bears a constant load and therefore remains of constant length. The other carries a micrometer screw and a pan for weights. A bubble-type level is hung between a point on the fixed wire and the top of the micrometer screw on the wire under test. The idea is to vary the stress on the test wire by adding weights to the pan and to measure the length change by adjusting the micrometer screw to bring the spirit level back to horizontal. There are two copies of the experiment; one uses stainless steel wire, the other ordinary carbon steel.

First two practical matters:

Firstly, note that one complete turn of the micrometer wheel produces a vertical movement of 1 mm. The wheel has 10 large divisions and each large division is subdivided 10 times. Consequently each of the subdivisions represents a vertical displacement of 10 μm or 0.01mm. Figure 3 shows a sample reading of 0.125mm. The bottom of the wheel is between 0 and 1mm as seen on the side mm scale. The reading on the wheel shows that it is beyond 0mm by 0.125 mm.

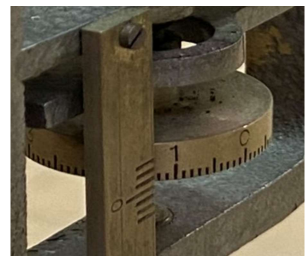


Figure 3: Sample reading on micrometer wheel of 0.125mm

Secondly, to initially remove kinks in the test wire, a 1 kg weight is added to the pan. This dead weight is left in the pan throughout the experiment and is otherwise ignored, i.e. it is not part of the measurements as such. Pre-straining the wire in this way takes it to a different part of the stress-strain curve, but because the relationship between stress and strain is linear in the elastic region, it does not matter exactly where you begin on the line.

1. With the 1 kg dead weight in the scale pan, adjust the micrometer screw until the level is horizontal. The observed reading serves as the origin from which extensions will be measured.

2. Place 100 g in the pan and bring the level back to the horizontal again with the micrometer. The difference in the micrometer readings is the extension of the wire for the 100 g load. Repeat with 200, 300 etc. up to 1000 g loads. Plot a graph of extension vs load as you obtain your data.

   Sometimes there is an observable curvature of the graph at lower loads but it becomes straight at high loads, especially with a new wire. This can be caused by residual kinks in the wire which have not completely straightened out under the initial load. If your graph shows such a behaviour with increasing load, try taking points up to a total loading of 2.5 kg (including the dead weight) if this is still within the range of the micrometer. All the higher points will fall on a straight line if the explanation is correct. If, however, the graph appears to be straight at lower loads and curves at high loads, the wire may be operating beyond its elastic limit and no extra load should be added. This is unlikely, but the two cases can be definitely distinguished by what happens when you decrease the load. In the former case the downward curve will be identical, at least at higher loads where kinks are not relevant. In the latter, the downward curve will follow a different path all the way down.

3. If the wire under test has not exceeded its elastic limit, it will return along the same curve to its original length when the load is removed. To test for this, measure the extension for loads of 950, 850 g, etc. as masses are removed from the scale pan. Use different symbols for increasing and decreasing loads in your graph.

5. Measure the diameter d of the wire in several places with a micrometer (Figure 4 shows how to read the micrometer), and measure its total length ℓ using a 2 m rule.

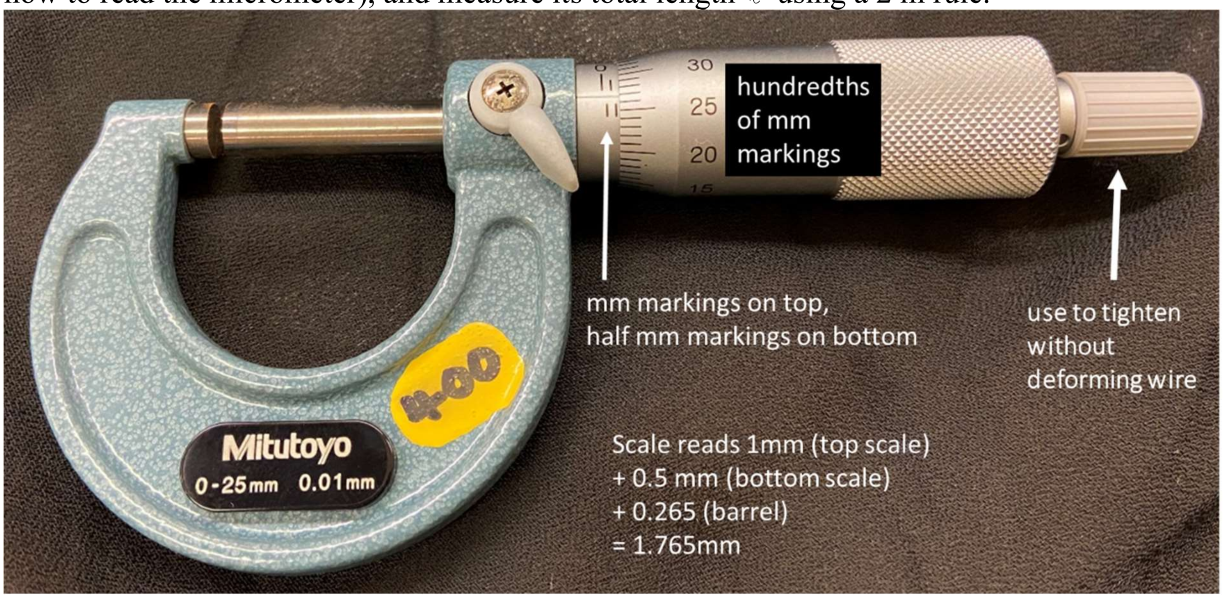


Figure 4: Sample reading of micrometer.

6. Check your value of E (if it isn't within 10% of the rough value of 200 GPa you should check your measurements).

(ii) Dynamic method

A long steel wire with a known mass already attached is available. You will need to measure the length and average diameter of the wire. The mass can be set into vertical oscillation by tapping it underneath with the rubber mallet. A semiconductor chip, which detects acceleration, is fastened to the mass and provides an output voltage signal proportional to $\ddot{x}$. This signal is examined on the oscilloscope. The sensitivity of the chip is 400 $mV/$g, where g is the acceleration due to gravity. (These chips are inexpensive and have interesting applications, e.g. in measuring acceleration during collisions or detecting vibrations, as here. There should be an information sheet and an application note available on the bench).

Comparing equations (3) and (4) it can be seen that the frequency of the sinusoidal variation in acceleration is the same as that for position thus the period of the acceleration signal seen on oscilloscope can be used to determine ω or f. E can then be calculated using Eq. 5. Also measure the initial amplitude of the sinusoidal signal so that you can determine the maximum strain and maximum stress reached.

Before you collect data set up the Tektronix TBS 2000 scope as follows:

Adjust Trigger

→ *Trigger Menu* → *Type = edge, Source = CH1 , Mode = Auto (otherwise the signal stops rolling)*

Check that scope is set for x1 probe.

→ *CH1* → *Probe Setup = 1X*

1. For a number of trials freeze the signal on the oscilloscope using the run/stop button. Do this soon after striking but once the oscillations become sinusoidal. Use the vertical timing lines (cursor function) to measure the time for a number of periods in the range of about 10 to 30 so that you can determine the frequency with minimal error. Use the horizontal cursors lines to measure the amplitude of the signal in mV. (*Use Multipurpose dial to move cursor, tap Multipurpose to switch between cursors.)*

   Overdriving the system with too hard a tap with the mallet will give unreliable data because the stress may go to zero when the mass is near the top of the oscillation and so the motion is no longer simple harmonic. It is also possible to break the wire if the wire exceeds its elastic limit, so take care. Do a quick calculation of E using one of the frequency values to make sure you are within 10% (otherwise you should remeasure the wire dimensions and frequency).

2. Another useful quantity to determine is the time taken for the oscillation to decay. This is a direct measure of the quality factor Q of the system which is given by $Q = 2\pi(total\ energy\ stored)/(energy\ lost\ per\ cycle)$. If you measure the time t for the amplitude to decay by a factor of one half, then Q is given by $Q = \pi tf / \ln 2$. (See *Mechanical Resonance*). It's interesting to ask how the energy is dissipated. Although there are losses due to viscous air damping and sound emission, there is also 'internal friction' in the wire itself. For several trials, enough to get good statistics, record the time for the amplitude to decay to one half.

Analysis and Report

1. For the static method
   a. Plot extension vs. load. Use linear regression to fit a straight line and obtain the slope of the graph for increasing loads. Repeat for decreasing loads. Examine the residuals (say $y_i - y_f$ where y_i are the extensions and y_f the values given by your fit) to look for any curvature in the graph. Are two lines identical within experimental error?
   b. If the lines appear to be the same and are straight, again use linear regression to obtain the slope of the graph using all available data points. If you detect any curvature, you must consider carefully how to proceed with your analysis, perhaps in consultation with an instructor.
   c. From the slope calculate E with an error estimate.
   d. Calculate the maximum strain and the maximum stress (don't forget to include the

initial 1kg dead weight).

2. For the dynamic method
    a. Evaluate E from your average value of f. Determine the uncertainty from the standard error on f.
    b. To determine the maximum strain and maximum stress reached you will need to determine the maximum acceleration and size of oscillations (in meters). The observed signal amplitude can be converted from mV to m/s^2 using the chip sensitivity (400 $mV/$ g). The amplitude of vibration x_0 can then be determined from equation (4). Using x_0 determine the maximum strain. Draw a free body diagram of the suspended mass and use the maximum acceleration to determine the maximum tension and stress in the wire.
    c. Determine the Q of the system.

3. Compare the values of E obtained by the two methods with the expected range of values. (The wires used are both carbon steel but may be different compositions). A **rough** value for steel is $E = 200$ GPa but check reference books/websites for the ranges appropriate for carbon steel. (e.g. *Tables of Physical and Chemical Constants*, Kaye and Laby, Longmans – link on course website, *MatWeb* - http://www.matweb.com/). Also look up the expected yield stress for steel. Were you close to this limit? Where your strain values less than the 0.1-1% discussed in the introduction?

Electrical Resonance

This experiment is an investigation of the response of a series combination of a resistor, capacitor and inductor when driven by an external oscillator.

Introduction

Electrical circuits may exhibit resonant behaviour when excited by an external oscillating voltage. One of the simplest examples is the RLC series circuit shown in Fig. 1. The external oscillator with angular frequency ω

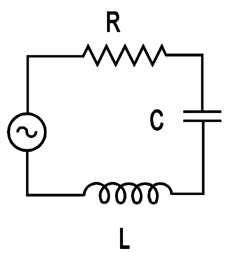


Fig. 1

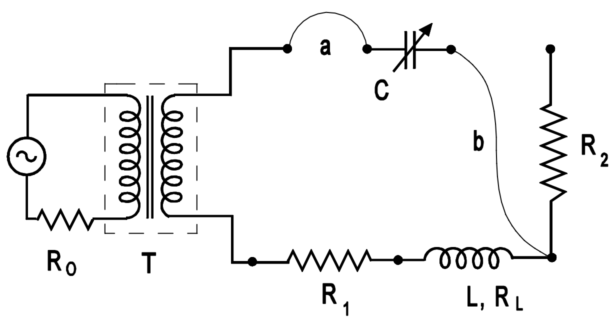


Fig. 2

produces current oscillations in the circuit at the same frequency. However, an LC circuit has a natural oscillation frequency[1] $\omega_o = \sqrt{1/LC}$ and when $\omega = \omega_o$ the current oscillations have a very large amplitude. This is called resonance. The phase of the current with respect to the applied voltage also undergoes large variations as the frequency of the driving voltage goes through resonance. Both the amplitude and the phase are investigated in this experiment.

Many mechanical systems also possess a natural oscillation frequency and exhibit resonance. Another experiment 'Mechanical Resonance' deals with an example of this type.

Theory

The sum of the voltages around the circuit of Fig. 1, in terms of the charge q on the capacitor, is

$$L\ddot{q} + R\dot{q} + \frac{q}{C} = V_o \cos \omega t \quad (1)$$

[1]When the series LC circuit has some resistance it is said to be damped. It then shows current oscillations with an exponentially decreasing amplitude when disturbed. The frequency of these oscillations is slightly shifted from $\omega_o = \sqrt{1/LC}$.

where $V_o \cos \omega t$ is the voltage at the output of the oscillator. The solution of a similar equation is dealt with in 'Mechanical Resonance'. The solution to this equation for the current, when all transients have died away, is found to be

$$I = \dot{q} = \frac{V_o \cos(\omega t - \phi)}{\sqrt{R^2 + (\omega L - \frac{1}{\omega C})^2}} \qquad (2)$$

where the phase angle ϕ is given by

$$\tan \phi = \left(\frac{\omega L - \frac{1}{\omega C}}{R} \right). \qquad (3)$$

The above results can also be obtained using complex number notation (see the experiment 'Electrical Impedance' for an introduction to this method). The current $\mathbf{I}$ is given by

$$\mathbf{I} = \frac{\mathbf{V}}{\mathbf{Z}} = \frac{V_o e^{j\omega t}}{R + j\omega L + (1/j\omega C)} \qquad (4)$$

The heavy type means the quantity has both amplitude and phase information. This equation immediately yields the desired results for the amplitude and phase response quoted above. More details may be found in Sec 10.6 and Appendix 10.1 in W. J. Duffin *Electricity and Magnetism*, or in B. I. Bleaney and B. Bleaney *Electricity and Magnetism* Ch 7 (or any other text on electromagnetic theory or electrical circuits).

Eq. 2 shows that, if ω is varied, the current is a maximum at $\omega_o = \sqrt{1/LC}$. The precise results obtained here are for the current in the circuit but one obtains similar results if, for example, the voltage across the capacitance is calculated. In this case Eqn. 4 is replaced by

$$\mathbf{V}_c = \mathbf{Z}_c \mathbf{I} = \frac{V_o e^{j\omega t}}{j\omega C(R + j\omega L + (1/j\omega C))} \qquad (5)$$

and note that the precise frequency observed for the maximum amplitude V_c as ω is varied is somewhat shifted from ω_o. The phase difference is also different, though the variation with frequency is the same except for a shift of $-\pi/2$.

There are various sources of stray capacitance in the actual circuit which affect the resonance frequency. There will be capacitance due to the coaxial cables to the oscilloscope and the input capacitance of the oscilloscope itself. However, these have little effect because they are in parallel with the small resistance R_1.

The main source is the inductor itself which has capacitance between the coils. This can be represented as a capacitance C_s across the inductance as in Fig. (3). Including this one finds that the resonance frequency is given by $\omega_o = \sqrt{1/L(C + C_s)}$. This can be written in a convenient form for the analysis in terms of the measured quantity $f_0 = \omega_0/2\pi$ as:

$$\frac{1}{f_o^2} = 4\pi^2 L(C + C_s) \qquad (6)$$

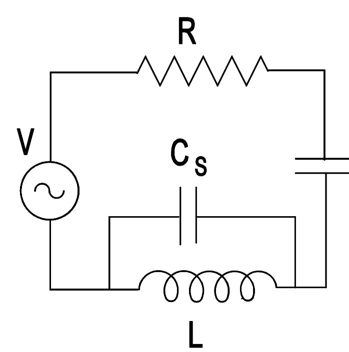


Fig. 3

Experiment

Fig. 2 is a sketch of the circuit board showing the circuit as used in the first part of the experiment with only R_1 connected; R_2 is used later. The inductance is fixed, but the capacitance can be varied. The oscillator provides a driving voltage of variable frequency and amplitude to the circuit via an isolation transformer. The transformer is needed to break the direct ground connection so that an oscilloscope can be used to measure potential differences anywhere in the RLC circuit. The output transformer acts as an AC voltage source in series with a somewhat uncertain impedance $\mathbf{Z}_T$. You can avoid having to consider $\mathbf{Z}_T$ *if you make sure to measure all oscillatory voltages and phases relative to the output of the transformer, say* $\mathbf{V}_T$. *In this case,* $\mathbf{V_T}$ *becomes exactly equivalent to* $V_o e^{j\omega t}$ *in Fig. 1.*

(1) Determine L and C_s by finding the resonance frequency f_o for various capacitance values C.

Set C ~ 3 nF to begin with. Connect the left input channel of the oscilloscope across the resistor R_1. This voltage is a direct measure of the current via $\mathbf{V_R} = R_1\mathbf{I}$ and contains both the amplitude and phase information. You can read the amplitude of the voltage V_R directly from the display as a function of frequency.

Scan through the frequency range to locate f_o. Record the values of f_o and C, and repeat for various values of C over a range of 1-10 nF in increments of 1 nF. Use Eq. (6) to plot an appropriate graph and deduce L and C_s.

(2) Examine amplitude and phase around resonance for $C = 3$ nF

On the newer scopes the amplitude and phase measurements (parts a and b) can be made at the same time using the Measure function. Keep the left hand channel across R_1, this is your circuit output and monitor the output of the transformer V_T using the right hand channel, this is the input to your circuit. Be careful to keep the grounds of the two channels on the same point.

Before you start, set up the following on the ***Tektronix*** *TBS 2000* scope

Adjust Trigger

→ Trigger Menu → Type = edge, Source = CH2 since the input to the circuit will be fairly consistent in size (no external trigger on this scope)

→ Trigger Level, adjust this knob until the signal is stable – the level is shown as a line on the screen and should be within the range of the signal i.e. not above or below it or too close to the top or bottom

→ Measure

→CH1 (circuit ouput i.e. R), Amplitude = Peak-to-Peak

→CH2 (input i.e. output of transformer), Amplitude = Peak-to-Peak, Time = Phase (Phase Ch2-Ch1)

Note: Keep the signals so that they are as large as possible on the screen but do not go off the screen, otherwise you will have faulty readings.

Check that both channels are set for x1 probe

→ CH1 → Probe Setup = 1X

→ CH2 → Probe Setup = 1X

During the experiment you should adjust the averaging used to eliminate noise as the output gets smaller (here is a video that explains this https://stream.queensu.ca/Watch/k7G8Nay5)

→ Horizontal Acquire →Mode= Average (rotate Multipurpose to move to this, press Multipurpose to select, rotate to change number of averages, push to select) – noisy signals require more averaging but this will slow down response to changes in the input. In this experiment you will need to adjust how much averaging you do. Away from resonance the signal gets smaller and the noise stays the same so you will need more averaging. Close to resonance the signal is large and changes rapidly so too much averaging will cause you to wait while the signal settles.

Do parts a) and b) in tandem.

a) Amplitude: Take enough readings of amplitude of V_R versus f to clearly define the curve in the neighbourhood of the resonance (around 25 points). Make sure you record both V_R and V_T because it is the ratio V_R/V_T that is important. Make an accurate measurement of the peak value of V_R/V_T which you will need in the analysis. Also take a few measurements to frequencies well above and well below resonance to determine the shape of the resonance over a wide frequency range. Remember to adjust the averaging to reduce noise when the output is small compared to the noise (i.e. away from resonance). A range of $f_o/2$ to $2f_o$ is suggested. Plot your points as you take them to make sure you have a well defined curve.

b) Phase: On new scopes the phase difference can be found using the Measure function as described above. The sign of the phase shift will depend on the order of the channels chosen. Plot ϕ as a function of frequency as you go to make sure rapidly changing sections of the graph are well defined.

(3) Compare voltages across components at resonance (magnitude and phase)

At f_o, use the oscilloscope to determine the ratio of the voltages across R_1 and L (use their midpoint as ground) and estimate the phase shift. Similarly, compare the voltages across L and C using their midpoint as ground. If L was ideal, the voltages across L and C would be expected to be equal in magnitude and opposite in phase (agreed?). While both channels have to have a common ground, consider the effect of using the midpoint between circuit elements as ground on the phase reported by the scope (ask if you are unsure).

(4) Investigate effect on the shape of the resonance curve of adding resistance

Increase the total resistance in the circuit by adding R_2 in series, then repeat the amplitude measurements (but probably take fewer points and keep to the region around the peak) and plot the new resonance curve on the same graph.

(5) Measure resistances so total resistance is known for theoretical curves

Measure the resistances R_1 and R_2, as well as that of the inductor R_L, with a multimeter. Disconnect all other circuit elements when you do this. Note that a DC measurement of resistance does not necessarily give exactly the AC resistance at high frequencies (look up 'skin-effect' in any book on electromagnetism to understand why), but the frequencies are low enough here that this should not be too important. A more serious effect will be hysteresis losses in the core of the inductance which will act like an extra resistance in the circuit. The resistance of the transformer is not relevant provided you always plot V_R/V_T.

Analysis

Plot a graph of $1/f_o^2$ versus C and use Eq. (6) to determine L and C_s.

Calculate the expected values of V_L/V_{R1} and V_L/V_c at resonance as well as the phase shifts and compare with the measured values in part (3). Notice V_c and V_L are much larger than the applied voltage V_T. Is this expected?

Compare the resonance curves for the circuits with and without R_2 by plotting V_R/V_T vs. f for the data collected in parts 2a and 4 on the same graph. Explain what happens when you add the extra resistor.

In order to make a detailed comparison of experiment with theory calculate the amplitude and the phase as a function of f (or ω). There will likely be some difference between the measured and predicted ω_o. This can be allowed for by making the comparison in reduced units, i.e. A/A_o (or ϕ) versus f/f_o or, ω/ω_o where A_o is the amplitude at the resonance frequency f_o. In this regard you can show that convenient forms for evaluating Eqs. (2) and (3) are

$$\frac{A}{A_o} = \left(1 + Q^2\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right)^2\right)^{-\frac{1}{2}} \quad (7)$$

$$\tan\phi = Q\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right) \quad (8)$$

where $Q = \omega_o L/R$ is known as the 'quality factor', and you can use $f/f_o = \omega/\omega_o$. Remember to include all the relevant resistance, e.g. $R = R_1 + R_L$ in part (2), but do not include the transformer resistance since this has been eliminated by the use of $A = V_R/V_T$. Evaluate the theoretical curves with lots of points (~100).

Now plot $\dfrac{V_R/V_T}{\left(V_R/V_T\right)_0} \; vs. \log(f/f_0)$ experimental points and theoretical curve (for circuit with only R_1) and ϕ *vs. log* (f/f_0) experimental points and theoretical curve.

Using a log scale for the f axis will show the data and calculations over a wide range and you should notice the symmetry of A/A_o and ϕ around f_o when the scale is of this form. You might need to adjust Q and/or f_0 to obtain the best fit of experiment to theory. This is best done using the amplitude results.

Mechanical Resonance

The resonant properties of a stretched wire driven by an external oscillator are investigated. This is an example of a forced oscillator.

Introduction

Many mechanical systems exhibit resonance phenomena at numerous frequencies. Examples are strings, air columns, membranes, quartz crystals in watches and computers, practically all solid objects and even the sun. The various modes of oscillation at these resonances are called the normal modes of vibration. In the case of the sun, the observation of the normal modes has provided a wealth of information about the sun's interior.

Another interesting recent development has been in what is now called micromachining. It is possible to make mechanical resonators which are only a few microns across. They can be incorporated into silicon microcircuits and, because they are so small, they can be made structurally perfect. This results in extremely high Q (quality) factors. Although most of the devices are relatively sophisticated, a simple tuning fork can be made which is only a few microns across with which it is possible to detect a vibration at a predetermined frequency. The development of micromachining is expected to be one of the major growth fields in the next decade.

This experiment examines the resonance of a simple stretched wire near its lowest resonance frequency. This system is the prototype of all stringed instruments.

Theory

The velocity of transverse waves on a stretched wire is $v = \sqrt{T/\sigma}$ where σ is the mass per unit length and T the tension. Resonance occurs when waves travelling in the two directions constructively interfere and satisfy the boundary conditions that the ends of the wire are fixed. This happens when $n\lambda/2 = L$, where λ is the wavelength and n an integer. (The appendix gives a more rigorous derivation). Thus the resonance frequencies are given by

$$f_n = \frac{v}{\lambda} = \frac{n}{2L}\sqrt{\frac{T}{\sigma}}. \qquad (1)$$

To find the response of the wire near one of the resonance frequencies, we model it as a damped harmonic oscillator which is driven by a force $F\cos\omega t$, *i.e.*

$$m\ddot{x} + b\dot{x} + kx = F\cos\omega t \qquad (2)$$

or $$\ddot{x} + \gamma\dot{x} + \omega_o^2 x = C\cos\omega t = \mathrm{RP}\,(Ce^{j\omega t}) \qquad (3)$$

where $\omega_o = 2\pi f_n$ is one of the resonant frequencies of the wire and $C = F/m$ (though F and m are not useful parameters for this experiment) and γ indicates the strength of the damping. The resonances of the wire are assumed to be far enough apart that they do not influence each other. The full solution of this equation corresponds to a transient part plus the steady state solution. To find the latter we assume it to be of the form $x = A\ \cos(\omega t + \phi) = \mathbf{A}e^{j\omega t}$ *where* $\mathbf{A}$ is a complex constant (giving both the amplitude A and phase ϕ of x) and substitute in Eq. (3) to find

$$\mathbf{A} = \frac{C}{(\omega_o^2 - \omega^2) + j\omega\gamma}. \qquad (4)$$

Thus the solution is

$$x = \frac{C\cos(\omega t + \phi)}{\sqrt{(\omega_o^2 - \omega^2)^2 + (\gamma\omega)^2}} \qquad (5)$$

where ϕ is given by $\tan\phi = -\gamma\omega/(\omega_o^2 - \omega^2)$, but is not measured in this experiment.

When ω_o is fixed, the amplitude A of the response varies with ω and is maximum, say A_o, when the denominator in Eqn. (5) is a minimum. By differentiation with respect to ω we find this to be at

$$\omega^2 = \omega_o^2 - \gamma^2/2. \qquad (6)$$

This is very close to $\omega = \omega_o$ if γ is not too large.[1]

It is often useful to define a 'quality factor' Q which tells us the sharpness of the resonance. There are various equivalent definitions of Q, one of which is $Q = \omega_o/\gamma$. Using this, the amplitude in Eq. (5) can be written in the form

$$A = \left(\frac{C}{\gamma\omega}\right)\left(\sqrt{\left(\frac{\omega_0}{\omega} - \frac{\omega}{\omega_o}\right)^2 Q^2 + 1}\right)^{-1} \qquad (7)$$

As above, if γ is small so that Q is large and the maximum amplitude A_o is close to $\omega = \omega_o$, then A_o is approximately $C/\gamma\omega_o$. The amplitude drops to $A_o/\sqrt{2}$ at $\omega_o \pm \Delta\omega$ where $\Delta\omega \approx \omega_o/2Q$, so that the width of the resonance at these points is just ω_o/Q.

The above theory assumes the oscillator is simple harmonic. This means that the restoring force increases linearly with displacement. Many oscillators are non-linear, and you may notice this in the present experiments. In the detailed theory it is assumed that T is a constant. In fact, as the wire displacement increases, T increases which leads to a restoring force increasing faster than the displacement, a situation referred to as a 'hard' oscillator. This is dealt with in various texts, e.g. H.J. Pain, *The Physics of Vibrations and Waves* 4th Edn, p. 394 or J.B. Marion, *Classical Dynamics*. The response is still periodic, but no longer precisely sinusoidal, i.e. simple harmonic. When such an oscillator is subjected to forced resonance, the resonance frequency depends on

[1] The transverse velocity $\dot{x}$ is found to have its maximum value at $\omega = \omega_o$, regardless of γ, and this behaviour carries over to electrical circuits where the current corresponds to velocity in the mechanical case. This is examined in the '*Electrical Resonance*' experiment.

the amplitude of the driving force increases. Also, the resonance peak becomes asymmetric about the maximum amplitude point. The resonance curve ‘leans’ toward higher frequencies for a hard oscillator. ‘Soft’ oscillators have all the opposite features.[2] If the driving force is increased, the amplitude becomes larger and the non-linearity becomes more prominent. As a final point, the amplitude of the response is no longer necessarily a linear function of the driving force amplitude. This is partly due to the resonance shifting with amplitude, but there is also a part due directly to the non-linear restoring force.

Experiment

A forced vibration of the wire is produced as follows. As sketched in Fig. 1 the non-magnetic wire hangs vertically in the gap of a permanent magnet. An electrical oscillator drives an ac current I through the wire, which results in an alternating force

$$\mathbf{F} = I\cos\omega t\int d\mathbf{l}\times\mathbf{B} \qquad (8)$$

in a horizontal direction perpendicular to the direction of the magnetic field $\mathbf{B}$. The integration is over the region where $\mathbf{B}$ is finite. The wire is illuminated and the amplitude of the forced oscillations may be obtained from the width of the band of light reflected from one edge of the wire using a travelling microscope. When you observe the vibrational motion of the wire you might notice that it moves in a circular or elliptical motion rather than linear. This is because the wire has two degenerate modes of vibration at right angles to each other, and energy fed into one of the modes usually leaks into the other. You are looking at both modes excited at the same time.

The lower end of the wire is soldered into a brass plug which is screwed into the middle of a heavy brass disk. To prevent the suspended system from swinging as a pendulum, an attached aluminum vane hanging between the poles of another permanent magnet provides electromagnetic damping. As the vane, disk, etc. are not ferromagnetic, no force is exerted on them by this magnet when they are at rest. (Even non-ferromagnetic materials feel some force due to their magnetism when placed in a magnetic field but this is negligibly small under typical conditions as here). To permit calculating the tension in the wire without dismantling the apparatus, the suspended weight is given.

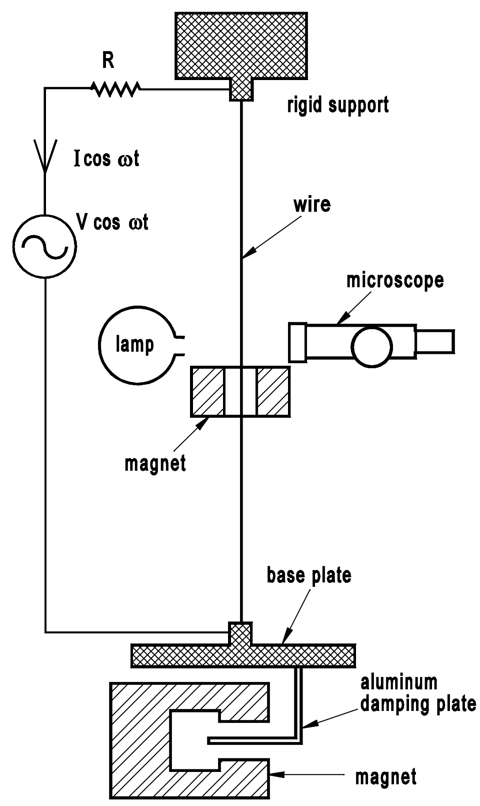


Fig. 1 Apparatus

[2] Eqs. (5) or (7) predict that a linear system also has an asymmetric response curve for sufficiently large variations in ω/ω_o. However, for a system of relatively high Q, the equation gives a curve departing negligibly from symmetry about the maximum amplitude. You will see this if you make a full comparison of your results with these equations.

(1) Setup

a. Switch on the viewing light and anything else that might cause heating and require time for stabilization. The room windows should be kept closed to prevent temperature changes.
b. Weigh the sample of stainless steel wire which. A sample of the same wire as the one you will use in the experiment should be available at the experimental station. Weigh the sample using a balance which gives mg resolution and measure the length so that you can calculate the mass per unit length σ.
c. Measure the length of the suspension wire L using a tape measure
d. Calculate the lowest resonance frequency f_1 using Eq. (1).
e. Measure the diameter of the wire; you will need it in the analysis.
f. Spend some time getting used to the oscillator controls. Ask for assistance if necessary. Note that the frequency output is accurate to 10ppm. Set the oscillator output to 2.5 V, leave the series resistance on maximum, and set the gain control on the power amplifier to 10-20% of full scale. The oscilloscope is available so that you can check that the waveform at the output of the amplifier is not distorted.

(2) f_1

Find the lowest frequency f_1 experimentally. When you are sure that you have resonance, set the amplifier gain control so that the vibration does not exceed the length of the microscope scale, or better, $80-90\%$ of it. This preliminary adjustment ensures that when a systematic series of readings is taken, the amplitude will not exceed full scale.

(3) Resonance curve around f_1

Take the temperature. Measure the amplitude of vibration A for frequencies f in the range of about ±1 Hz around the resonant frequency f. Note that the oscillator has an extremely accurate and stable frequency output (~1 in10^5) because it synthesizes the output waveform using a crystal controlled oscillator as the reference. The frequency must be changed in quite small steps near resonance, probably in increments of 0.01 Hz, because the amplitude will vary strongly in this region. Near the ends of the range increments of ~ 0.05 – 0.1 Hz will be adequate. After each change in frequency, time must be allowed for the transient motion to die out. Qualitatively note the variations in amplitude which occur following the change of f_1, and how the period of this variation changes with deviations from resonance. Theory shows that the transient variations in amplitude represent beats at ~ $|f - f_1|$ between the forced oscillation frequency f and the natural, damped oscillation of the system at approximately f_1. Thus when you are close to resonance, the amplitude might take a long time to settle to its final value. At the end, check the temperature and the reproducibility of a few of the readings.

(4) f_3

Look for the resonance at f_3. Determine the resonance frequency and amplitude at that point, but do not attempt to determine the full curve. The amplitude will be lower, but do not adjust the settings of the oscillator voltage or the amplifier.

(5) Amplitude of the oscillation as a function of drive voltage V

With the frequency set for resonance as found in (2) measure the amplitude as a function of oscillator output voltage. Proceed as follows. Without touching the amplifier gain or oscillator voltage (still set to 2.5 V), first check that the resonant frequency has not drifted. If it has, reset the frequency to give maximum amplitude. Also check that there is no visible distortion from a sinusoidal waveform of the amplifier output. Now measure the vibration amplitude as you decrease the voltage incrementally down to 0.1 V. When you have finished, set the voltage to 1.0 V and determine if there is any visible shift of the resonance frequency compared to the value you found at 2.5 V drive.

Analysis and Report

(i) Calculate the resonance frequencies according to Eq. (1) and compare with the experimental frequencies. Is $f_3 = 3f_1$ as expected? The following lists various corrections to the resonance frequencies that may be significant. Estimate their magnitude for f_1.

- Eq. (6) shows that the frequency observed at maximum amplitude is shifted from f_1 by damping effects. Since $Q = \omega_1/\gamma$, we can readily estimate this correction. (Estimate Q from (iii) below for this).

- The value used for σ was for the unstretched wire. Estimate the correction due to stretching using the wire diameter and Young's modulus.

- In deriving the expression for f_n in Eq. (1) it was assumed that the wire is perfectly flexible, that the ends are immovable, and that none of the surrounding medium takes part in the motion. None of these are exactly true. The problem of the stiffness of the wire was treated by Lord Rayleigh, '*Theory of Sound*', p. 207 and pp. 300-301, Vol. I. Assuming both ends of the wire may be regarded as clamped, the appropriate correction is given in Eqn. (8) on p. 300. This may be expressed in the form

$$f = f_1(1+p) \quad \text{where} \quad p = \frac{r}{2L}\sqrt{(1+AE/T)} \qquad (9)$$

In these equations, r, A and E are the radius, cross-sectional area and Young's modulus for the wire.

(ii) The most striking characteristic of the forced oscillator with small damping is the large response at the resonant frequency and the steep decrease in response as the applied frequency deviates from the resonant value. Plot the resonance curve around f_1 using reduced units for the amplitude, i.e. A/A_o as a function of f.

(iii) Find the frequencies for which the amplitude A is $1/\sqrt{2}$ of the maximum amplitude A_o, and calculate the quality factor Q (see paragraph after Eq. (7)).

(iv) It is always preferable to make as detailed a comparison of experiment and theory as possible. In this case, use the value of Q found in (iii) and plot the theoretical resonance curve for A/A_o around f_1 as given by Eq. (7) on the same graph as your experimental data. Do this by evaluating A/A_o at closely spaced intervals (~ 100?) and draw a curve through these theoretical points (but do not use any symbols for these points). How well does this full theory agree with the experiments? You might need to adjust Q, and perhaps f_1 for the best fit.

(v) Plot the amplitude of the oscillation as a function of drive voltage V. The driving force on the wire is, of course, proportional to V, and a linear oscillator would give a straight line graph.

(vi) You may observe departures from the predicted behaviour due to non-linearities in the system. These might include: a leaning of the resonance curve, a non-linear response of the amplitude to the driving force, and a shifting of the resonance at high amplitudes. Are any of these effects visible?

(vii) You cannot excite the resonances at even multiples of the fundamental because of the placing of the magnet providing the exciting force. Explain why this is so.

Appendix: Normal modes

The way in which the system vibrates is different for each 'normal mode'. Each mode has a specific frequency of oscillation and a unique standing wave pattern and is a solution of the wave equation, which we write in 1D as

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 y}{\partial t^2}, \qquad (10)$$

with the boundary conditions appropriate to the problem. Here, y is the displacement of the particle at x. In 1D the standing wave solutions are of the form

$$y = f(x)\cos\omega t \qquad (11)$$

where $f(x)$ is the 'pattern' of the particular normal mode and ω its frequency. If this is substituted into Eq. (10), we obtain

$$\frac{d^2 f}{dx^2} = -\frac{\omega^2}{v^2} f = -k^2 f \qquad (12)$$

Here we use the result $v = \omega k$ where $k = 2\pi/\lambda$. We have dropped the partial derivative notation since $f(x)$ is only a function of x. This equation is equivalent to the simple harmonic oscillator equation, and has a solution

$$f(x) = a\sin kx + b\cos kx \qquad (13)$$

where a and b are two constants that must be obtained from the boundary conditions.

In the present case of a stretched string, let's put one end of the string at $x = 0$ and the other at $x = L$. Since $y = 0$ at the fixed ends, then substituting $y = 0$ at $x = 0$ gives $a\sin 0 + b\cos 0 = 0$, or $b = 0$. Now using $y = 0$ at $x = L$ gives $a\sin kL = 0$, or $L = n\lambda/2$. This is the result quoted in the text.

Eqs. (10) and (11) apply to other continuous systems. For example, organ pipes or the *Speed of Sound* experiment using resonance in a tube in first term are solved in a similar manner.

Interference and Diffraction of Ultrasound

The interference and diffraction of ultrasonic waves by narrow slits is investigated.

Introduction

For everyday purposes it is often assumed waves (light, sound, …) propagate in straight lines. However, if a beam meets an obstacle, a close examination shows that some of the beam penetrates into the region of the *geometric shadow*. This spreading is referred to as *diffraction*. It is also found that beams which take various paths from the source to the detector interfere with each other giving *interference* effects. Interference and diffraction are fundamental to physics. On a large scale their effects are observed in musical instruments (acoustic or vibrational waves both inside the instrument and the room), in the propagation of electromagnetic radiation (e.g. radio and microwave), in the fundamental limitations of lenses (in particular in telescopes and microscopes), in the propagation of light in optical fibres and on IC chips, and in the strange behaviour of matter on an atomic scale, etc., etc. A useful introduction to interference and diffraction can be found in the first year text, and the topic will be studied in detail in ENPH 211.

The present experiment will demonstrate these effects using an obstacle that has one or more equi-spaced parallel slits. Young's original 2-slit experiment is used as a starting point, and then the investigation is extended to examine happens with a single slit and with many slits. When the obstacle has many slits it is usually called a *diffraction grating*. Ultrasonic waves, that is acoustic waves that have a frequency higher than the human ear can detect, will be used. Sound has the advantage that the waves are of long wavelength (~ 1 cm in our case) so that a large-scale apparatus can be used. Ultrasound makes the experiment appear to be quiet, though it is very noisy.

Theory

When examining a single narrow slit acting alone, it is found that the emerging wave diffracts (spreads) over a wide angle, with a relatively uniform intensity as a function of angle. With more (narrow) slits opened, the emerging beams interfere with each other giving a series of low and high intensity bands parallel to the slits. The positions of the maxima for this case can be readily derived.

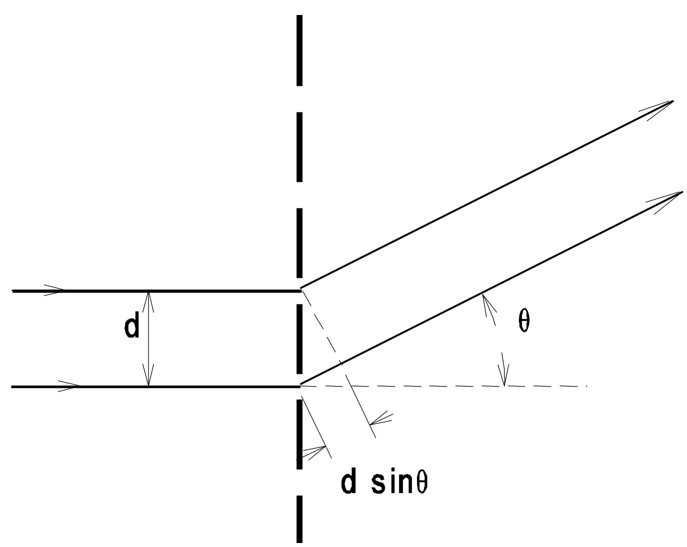


Figure 1. A diffraction grating

The simplest geometry, and the one assumed throughout, is with the receiver at a large distance from the slits (compared to their spacing). Emerging waves from each slit are assumed to still diffract over large angles with many slits open, just as with a single slit. In Fig. 1 parallel beams from two adjacent slits are shown at angle θ (angles are always measured from the normal of the grating). Note that beams which do not travel at the same angle will not interfere as the separation distance between transmitter and receiver approaches infinity. When the path difference between the beams is an integer multiple of the wavelength λ then they will interfere constructively and a maximum is expected, i.e.,

$$n\lambda = d\sin\theta \qquad \text{(for maxima with n = 0, 1, …)} \qquad (1)$$

where d is the spacing between the slits. If this is true for adjacent beams, it is also true for all beams. Hence this is a general result for finding maxima. This is a valuable result, but it does not tell us what happens at other angles.

Consider two emerging beams travelling parallel to each other as in Fig. 2. The beams are taken to be travelling waves of the form $A_\circ \cos(kx - \omega t)$ where ω is the angular frequency and $k = 2\pi/\lambda$ is the wave number, λ being the wavelength. The quantity $(kx - \omega t)$ is called the *phase* α of the wave. The superposition of the travelling waves from the two slits is calculated,

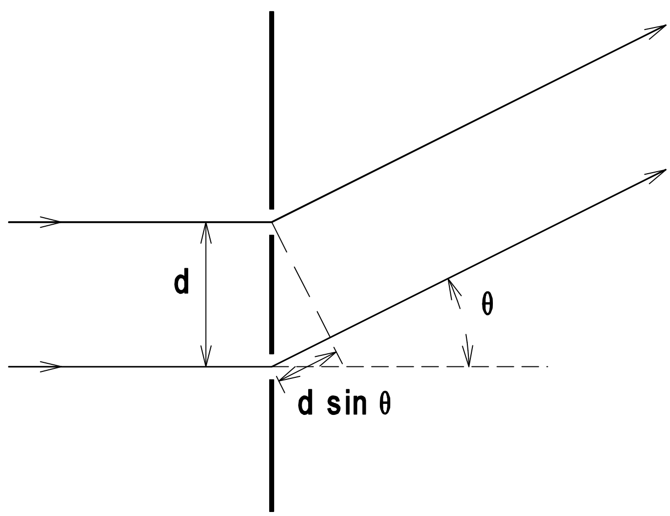


Figure 2. Two narrow slits.

$$\psi = A_\circ \cos(kx_1 - \omega t) + A_\circ \cos(kx_2 - \omega t) \qquad (2)$$

where $\psi = A\cos(kx - \omega t)$ is the resultant wave and the waves are assumed to have the same amplitude $A_\circ$. The two beams travel different distances x_1 and x_2 from the slits to the detector.

At the detector the time t is the same for both beams, but because the x_i are generally different their phases are also different. Eq. 2 and others of a similar nature are most easily evaluated using complex numbers, and so are written as,

$$Ae^{j\alpha} = A_\circ e^{j\alpha_1} + A_\circ e^{j\alpha_2} \qquad (3)$$

where it is assumed that only the real parts of the complex numbers are relevant. The amplitudes $A_\circ$ will actually not be a simple constant but will be a function of distance travelled because the wavefront is spreading, but this is ignored.

To find the amplitude A of the resultant the identity $A^2 = Ae^{j\alpha} \times c.c.$ is used where $c.c.$ means the complex conjugate (in this case $Ae^{-j\alpha}$). Thus $A^2 = (A_\circ e^{j\alpha_1} + A_\circ e^{j\alpha_2}) \times c.c.$ After some manipulation this can be written,

$$A^2 = A_\circ^2(2 + e^{\beta}) = 4A_\circ^2 \cos^2(\beta/2) \qquad \text{(2 narrow slits)} \qquad (4)$$

where $\alpha_1 - \alpha_2 = k(x_1 - x_2) = (2\pi/\lambda)d\sin\theta = \beta$. Notice that this equation predicts the maxima correctly at $\beta/2 = n\pi$ which is equivalent to Eq. 1 and it also tells us what happens at all angles.

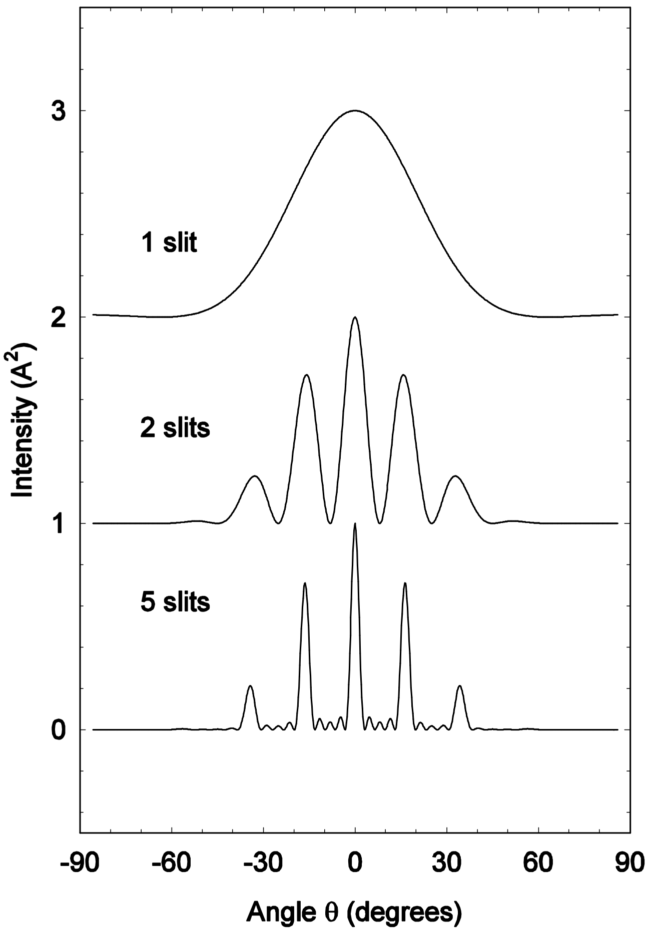


Figure 3. Calculated results. Each curve has been normalised so the central peak has unit intensity and also shifted vertically for clarity.

The receiver (microphone) actually responds to the amplitude A of the wave in these experiments but it is usual to deal with intensity, $I \propto A^2$ (as in the above equation) rather than amplitude. This is why all the equations are left in terms of A^2. In experiments with light the detectors typically respond to the intensity I not A.

Another result that is needed is what happens with a single slit of arbitrary width. This case is examined in the Appendix where it is shown that,

$$A^2 = A_1^2 \frac{\sin^2(\phi/2)}{(\phi/2)^2} \qquad \text{(single slit)} \qquad (5)$$

where A_1 is a constant and ϕ is given by $\phi = kb\sin\theta$. This equation is plotted as the top trace in Fig. 3. Notice that the form of the curve is due to interference of the beams from the various parts of the slit.

One could also use the same method to derive what happens with a double slit of spacing d and each of finite width b. This will not be pursued in detail here but the result is,

$$A^2 = A_2^2 \frac{\sin^2(\phi/2)}{(\phi/2)^2} \times \cos^2(\beta/2) \qquad \text{(2 slits of arbitrary width)} \qquad (6)$$

where A_2 is another constant. This is just the product of the single slit result, Eq. 5, and the (narrow) double slit result, Eq. 4. This result is shown as the middle curve in Fig. 3.

Finally, the general result for N slits is given as,

$$A^2 = A_N^2 \frac{\sin^2(\phi/2)}{(\phi/2)^2} \times \frac{\sin^2(N\beta/2)}{\sin^2(\beta/2)} \qquad \text{(general result)} \qquad (7)$$

where A_N is another constant. Again the single slit result is evident in the answer.

This equation reduces to Eq. 6 when $N = 2$, and Eq. 5 when $N = 1$. In other words, this is a general result which covers all cases. An example with $N = 5$ is shown in the bottom curve of Fig. 3.

In Fig. 3, notice that as the number of slits increases from 2 to 5, the peaks become sharper (and they continue to become sharper as N increases). However the peak positions remain unchanged and are still given correctly by Eq. 1. In Eq. 7 the peak positions are when the factor $\sin^2(\beta/2) = 0$ in the denominator.

Experiment

The experiment uses two ultrasonic transducers,[1] one operating as a transmitter at about 40 kHz and the other as a detector or receiver. These transducers are based on piezoelectric materials which change their length slightly when subject to an electric field. To make a transmitter, an oscillating voltage is applied to opposite sides of a piezoelectric thin disc which then undergoes an oscillatory change in thickness, thus launching an acoustic wave. These work best when the mechanical induced vibration is at a resonance frequency of the disc. The piezoelectric effect also works in reverse, i.e. changing the thickness of the disc produces an electric field which can be detected as a voltage difference across the surface. This gives a detector of sound waves.

The transmitter is placed at the focus of a large parabolic mirror to produce plane waves which are then incident on a metal grating with a number of equi-spaced slits. Two blank screens are used to mask some of the slits so that we can examine interference effects for an arbitrary number of slits. The whole assembly of transmitter, mirror and slits is supported on a turntable which is driven by the DC supply via a reversing switch so that rotations in either direction are possible. There is a 0-10 $k\ \Omega$ resistance potentiometer which rotates with the turntable and the varying potential is measured by a digital voltmeter (to get a measure of the angle through which

[1]An ultrasonic transducer is a device which converts a mechanical displacement into an electrical signal, or vice-versa.

the turntable has rotated). There should be another potentiometer on the bench that you can examine if you have not seen one before.

The receiver is held by a fixed stand at the same height as the transmitter and is connected to a variable-gain amplifier. The AC output of the amplifier is examined with an oscilloscope, and its amplitude is measured with the digital voltmeter.

Procedure

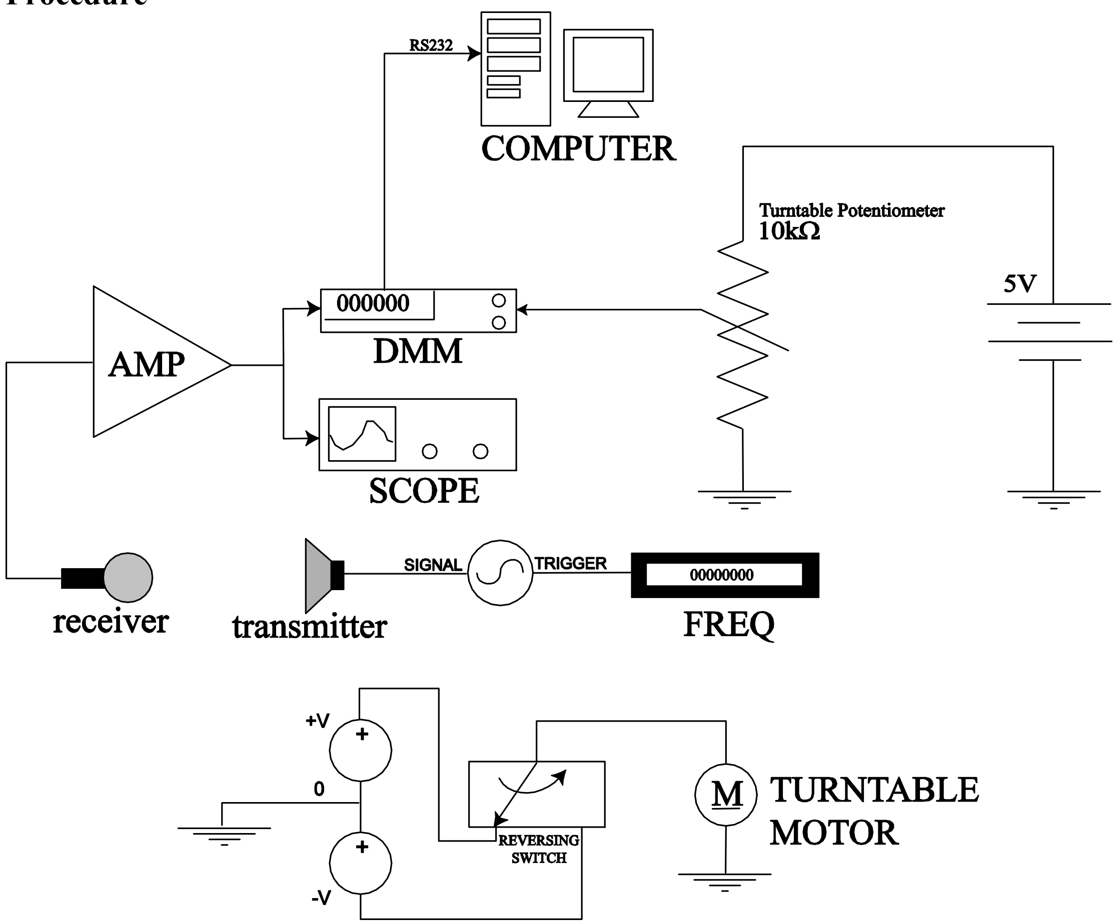


Figure 4. Apparatus Schematic

A: Set up and adjustments

1. Assure the apparatus is setup according to the schematic of Fig. 4.

2. Make sure the receiver and transmitter are placed well apart so that the assumption of infinite separation is well justified. Record this distance. Turn on all the equipment on the table and adjust the diffraction grating on the transmitter so that six slits are open (the screens should go to the edges of the open slits).
Note: Make sure that you hold the switch on the amplifier and signal generator in the on position for a few seconds. There is an automatic cutoff that may cut in if the switch is not held for a few seconds. This cutoff is to conserve the battery in the devices.

3. Set the DC level of the variable power supply to approximately 2.5V. This level will produce slow rotations of the turntable so that fine-tuning of the position is possible while still supplying enough voltage to start the motor up reliably. Adjust the turntable so that it is oriented perpendicular to the line between the receiver and transmitter. This position will roughly correspond to an interference maximum.

4. While observing the signal on the oscilloscope, adjust the angle using the motor switch until the largest amplitude signal is observed. This will be the point where the wavefronts are exactly perpendicular to the line of propagation from receiver to transmitter. Record the DC voltage measured on the digital multimeter (DMM), this will be the $\theta = 0^\circ$ position.

5. Keeping the turntable adjusted to the $\theta = 0^\circ$ position adjust the oscillator frequency until a maximum level is observed on the oscilloscope again. This will match the resonance frequency of the receiver with the frequency of the transmitter. Ensure that the amplifier is not clipping (see Figure 5) the signal while determining this maximum point. When the resonance frequency has been determined make note of it and then proceed to adjust the amplifier so that it is at the largest possible amplification before clipping occurs. It is important that clipping is never observed on the oscilloscope throughout the duration of this experiment.

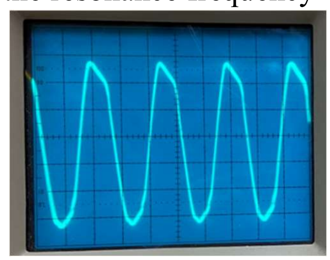


Fig. 5 Oscilloscope trace with clipping, Signal is no longer sinusoidal

B: Collect amplitude as a function of angle for 6, 4, 2, 1 and 0 slits open

6. Rotate the transmitter turntable so that it is offset about forty-five degrees from the perpendicular position. This will correspond to approximately 0.7VDC offset on the multimeter. Start the 'Interference and Diffraction' software on the computer and then press the *Collect Point* button. This will record and plot the AC RMS voltage being received at the detector against the DC voltage corresponding to the angle of the transmitter. Using the motor controller, rotate the turntable a small angle and then collect another point. Collect points covering a sweep of approximately 90° - your sweep should include the central peak and the two secondary peaks on each side of the central peak. Points should be recorded more frequently near maxima to make sure the peaks are not cut off.

7. When you are satisfied with the pattern the computer is displaying save the data points to a USB key for later analysis. It is recommended that the file is examined in Excel before clearing out the data points to ensure that all the data has been properly recorded.

8. Change the number of open slits on the transmitter to four and take a new sweep of the interference pattern as in step 6.

9. Repeat the process again for two open slits and the single open slit. Fewer data points are required as the number of slits are decreased since the interference and diffraction pattern will not change as rapidly. See Fig. 3.

10. You will likely observe significant distortion to the single slit data that you have obtained in comparison to the theoretical plots of Fig. 3. Close all the slits and take a measurement of the noise using another sweep. Since all the slits are closed, ideally, you would expect no signal to be observed. What is measured will represent the noise in the system. There will be both systematic and background noise present.

11. A conversion factor for the DC voltages measured at the turntable to angles between the receiver and transmitter is needed. This can be found by recording the voltage for several known angles and plotting angle vs. voltage. If the potentiometer is linear then the slope of the line can be used to determine the conversion factor.

C: Measurement of slit spacing *d*, width *b* (these affect peak position and relative height)

12. Take measurements of the width of the slits *b* and the width of the slats between slits, say *s*. In Fig. 1 it can be seen that the centre to centre spacing between slits $d = b/2 + s + b/2$ or simply $d = b+s$. You will find that the theoretical plots will be critically depend on the accuracy and uncertainty in these values. Make sure to record uncertainties in *b* and *s*. Figure 6 explains how to read the calipers.

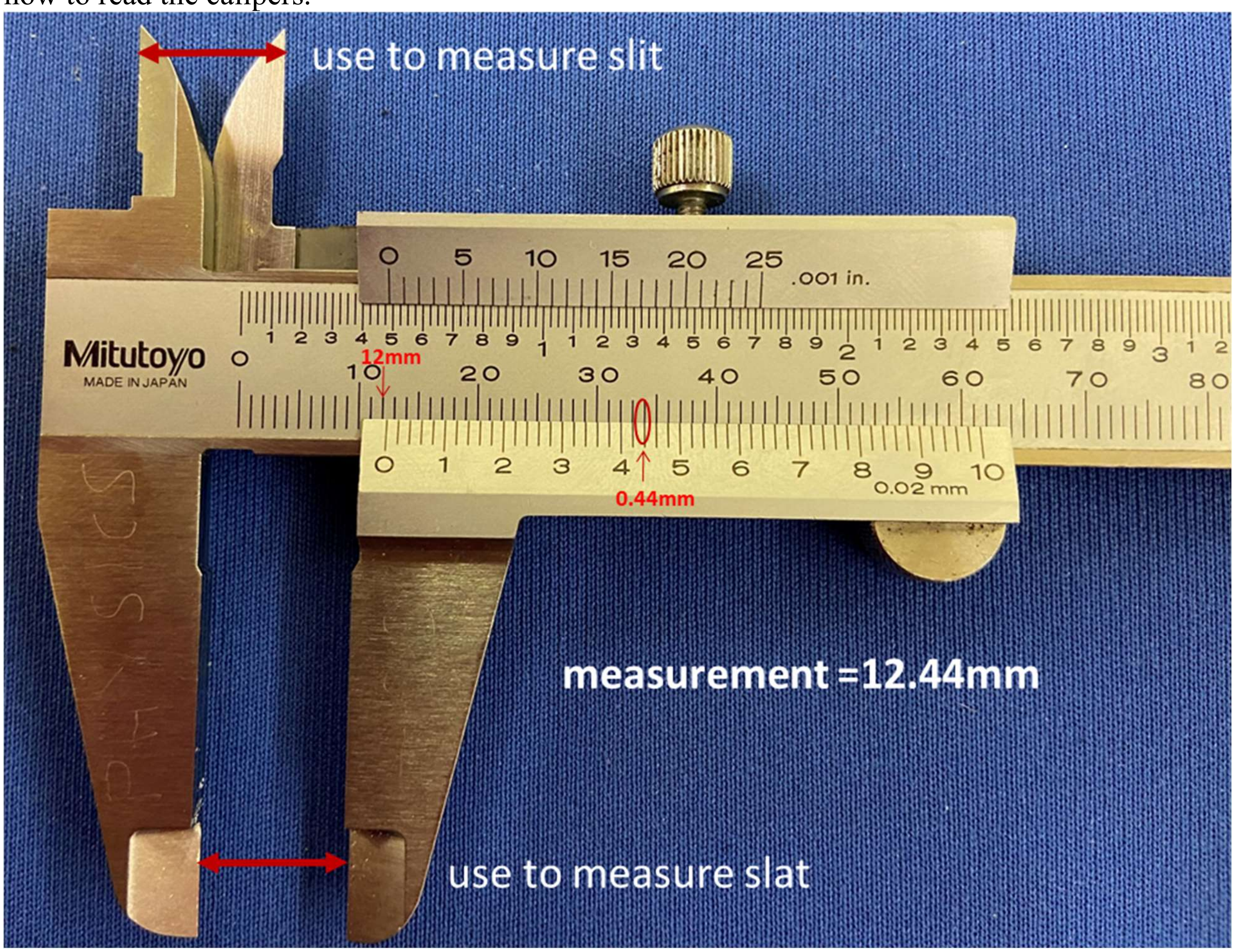


*Fig. 6 The calipers are read by first looking at the 0 on the bottom scale, in this case it is just beyond 12mm, then look at which line on the bottom scale lines up with one of the lines on the scale above, in this case it is two lines after the 4 which means 4 tenths + 0.02*2 or 0.44mm. The*

final reading then is 12.44mm. This is the distance between the jaws at the bottom and also the distance between the points at the top.

D: Measurement of air temperature to determine wave speed

13. Record the ambient air temperature using the thermometer supplied on the wall near the lab bench. You will need this to determine the speed of sound for your wavelength calculation.

Analysis and Report

1. The multimeter makes readings of both the turntable position and the amplitude of the signal received. Explain how it is possible for these signals to be distinguished.

2. Convert the potentiometer voltages to angles with $\Theta=0$ corresponding to the location of the central peak. Plot graphs of A^2 versus Θ for each data set. Ensure that the plot is vertically offset an appropriate amount to bring the data to zero.

3. Theoretical comparison:

   a. Note that the value of A_N in equation 7 is not known and so must be estimated from the experimental data. The maximum value of the other terms in Eq. 7 is N^2 so the theoretical central peak height will match the experimental central peak height if you set A_N= (height of experimental central peak)/N. Calculate the A_N value for each N.
   b. Use Eq. 7 to calculate the expected amplitude^2 versus angle for each data set (cf. Fig. 3) and plot the curves directly on the graphs of experimental data. As usual, use lines only for the theoretical curves.
   c. Include error bars only for the secondary peaks. Calculate the uncertainties in expected peak positions using Eq. 1. Using Eq. 7, and δb calculate the uncertainty in expected peak amplitude for the secondary peaks of each data set.
   d. Discuss the agreement and/or differences between the experimental and theoretical curves.

N.B. In calculating these curves, note that factors like $\sin(nx)/\sin x$ or $\sin(nx)/x$ appear, with n an integer. At $x=0$ or $\sin x=0$ the denominator is zero which will give divide-by-zero problems in the computations. However, simply avoid these points in the calculation and note that the results here can readily be calculated using $\sin(nx)\rightarrow nx$ for small enough x.

4. Plot A versus Θ for both the single slit data and the noise data on the same graph. Do you observe any correlation between the noise and the single slit data? If so, propose reasons that may explain this correlation.

Appendix

This gives the details of how Eq. 5 for a single slit is obtained. Note that this pattern is still a result of interference. The calculation treats the problem as there were a large number of narrow slits, each of width Δy, placed side by side with no obstacles in between (see Fig. 5). The amplitude of the wave at the receiver from each of these elementary slits is proportional to the width Δy.

The resultant from all the (parallel) beams at some angle θ is then,

$$Ae^{j\alpha} = C\sum \Delta y e^{jky\sin\theta} \qquad (8)$$

where C is a constant, $ky\sin\theta$ is the phase shift between the beam at the edge of the slit at $y = 0$ and the beam at distance y, and the sum is over all beams, i.e., all the elementary slits.

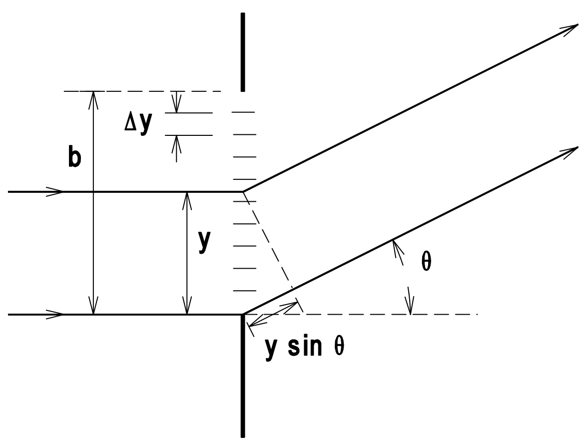


Figure 5. Single slit. Two beams are shown, one just grazing the lower edge of the slit, the other at a distance y ***above***

Taking the integral as the limit of the sum when Δy becomes infinitely narrow,

$$Ae^{j\alpha} = C\int_0^b dy\, e^{jky\sin\theta} = \frac{Cb}{j\phi}(e^{j\phi} - 1) \qquad (9)$$

where $\phi = kb\sin\theta$. Again using $A^2 = Ae^{j\alpha} \times c.c$,

$$A^2 = \left(\frac{Cb}{\phi}\right)^2 \sin^2\left(\frac{\phi}{2}\right) = A_1^2 \frac{\sin^2(\phi/2)}{(\phi/2)^2} \qquad (10)$$

where $A_1 = Cb/2$ is an amplitude constant.

A. P. French *Vibrations and Waves* (Norton).
R. S. Longhurst *Geometrical and Physical Optics* (Longmans)

The e/m ratio

The electronic charge-to-mass ratio e/m is measured by injecting electrons into a magnetic field and examining their trajectories. The data collected will also be used to estimate the magnitude of the earth's magnetic field.

Introduction

In this experiment we examine the trajectories of electrons in a magnetic field. Electron deflection by magnetic fields is used in many devices and instruments, including video monitors. This experiment also enables a measurement of the electronic charge-to-mass (e/m) ratio.

Consider an electron accelerated through a potential difference V. The electron gains kinetic energy of eV. Consequently

$$\frac{1}{2}mv^2 = eV. \qquad (1)$$

The electron is now injected into a uniform magnetic field B such that the direction of motion is perpendicular to the field lines. We can equate the force on the electron $e\vec{v} \times \vec{B}$ to the (mass × centripetal acceleration) to obtain

$$evB = \frac{mv^2}{r} \qquad (2)$$

where r is the radius of the orbit. Combining this with Eq.(1) we have an expression for the charge to mass ratio

$$\frac{e}{m} = \frac{2V}{B^2 r^2}. \qquad (3)$$

The experiment involves injecting electrons into a magnetic field that is generated by a pair of coils. The main difference between the experiment described above and the one you will perform is that the magnetic field will not be spatially uniform. It is very difficult to generate a spatially uniform magnetic field so we will have to make allowances for the non-uniformity in our calculations.

Experiment

The electron trajectories are measured in a vacuum tube which contains hydrogen gas at a very low pressure. Some of the electrons collide with the hydrogen gas molecules. The ionized gas emits a faint blue light which can easily be seen with the lights off. The pressure of the hydrogen gas in the tube ($\approx 10^{-2}$ mm Hg) is selected so that the space charge produced by the electrons bunches the beam for this range of velocities.

The electrons are generated by heating a filament. This gives some of the electrons in the filament enough energy to escape from the metal. These electrons are then accelerated by a potential difference applied between the filament (the cathode) and the metal cover surrounding the filament (the anode). After they have escaped through the nose cone of the anode they can be deflected and focused by a grid which is oriented parallel to the beam. We refer to the current in the filament as the filament current, the voltage between the filament and the metal cover as the anode voltage, and the voltage between the deflector grids as the grid bias. After the electrostatic field, the electrons encounter a magnetic field. The magnetic field is generated by a pair of coils which are placed above and below the tube in approximately a Helmholtz configuration, producing a fairly uniform magnetic field over a considerable portion of the volume between them. The ideal Helmholtz configuration has the circular coils separated by their radius R but our coils are not exactly in this configuration. The magnetic field is maximum midway between the coils on axis. For a pair with radius R and spacing $2b$, the magnetic field at the center of the coils is

$$B_o = \frac{\mu_o n I R^2}{(R^2 + b^2)^{3/2}} \qquad (4)$$

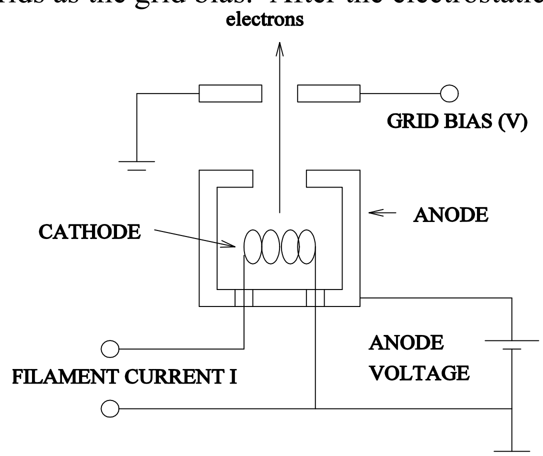


Figure 1. Schematic of the cathode and the electron optics used to accelerate the electrons.

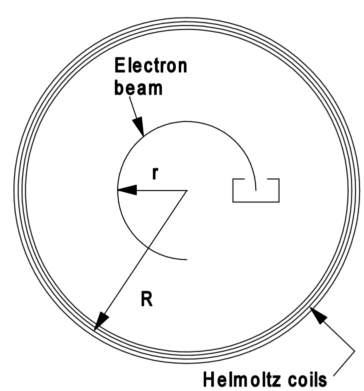
where n is the number of turns. In the Helmoltz configuration, $2b = R$ and so $B_o = 8\mu_o nI/5^{3/2} R$. We will write Eq. (4) as

$$B_o = KI \qquad (5)$$

where K is a constant for a given pair of coils and B_o is measured in units of Tesla. The pair that you will use have R = 15.4 ± 0.5 cm, 2b = 15.0 ± 0.5 cm and n = 130 turns. They generate a flux density which is measured to be KI with $K = (7.73 \pm 0.04) \times 10^{-4}$ T/A.

The magnitude of the field at the central point between the coils is of limited usefulness since the electron traverses an orbit in a plane parallel to the plane of the coils but at a distance r from the centre. By inspection the field is radially symmetric and we can allow for the fact that the field will be smaller as we move out from the centre.

Fig.2 Plan view of the electron trajectory which is approximately a circle with radius r, and the Helmholtz coils which have radius R.



To estimate the field at a distance r from the centre use the following data which was supplied with the coils. The normalized distance from the centre point r/R is listed with the normalized field B/B_o, where B_o is the field at the centre. This correction is shown graphically in Figure 3.

r/R	B/B₀
0	1
0.1	0.99996
0.2	0.99928
0.3	0.99621
0.4	0.98728
0.5	0.96663
0.6	0.92525
0.7	0.85121
0.8	0.73324
0.9	0.56991
1	0.38007

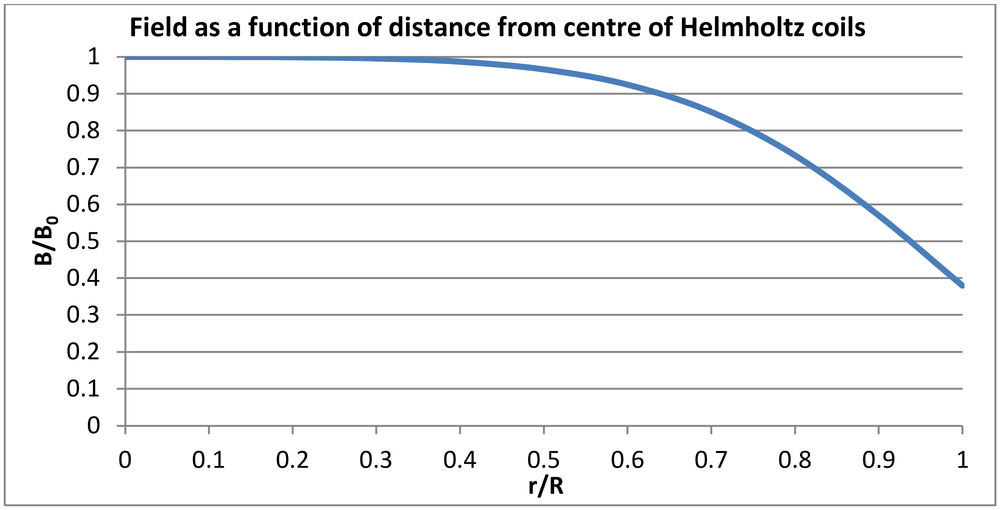


Figure 3 Normalized field as a function of normalized distance from centre of Helmohotz coils.

This adjustment can be incorporated into the constant K to produce a value K_r specific to the distance of the electron beam from the central point (i.e. radius of the orbit of the beam):

$$K_r = \frac{B}{B_0} K \qquad (6)$$

thus the field produced by the coils at a distance r from the central point is $B = K_r I$.

The magnetic field produced by the coils is not the only one present. The magnetic field of the earth, B_E ,will contribute to the total magnetic field, B_T ,experienced by the charges. In order to take this into account the coils are mounted so that the earth's magnetic field is parallel to the field from the coils. The tube can be rotated so that the field from the coils is in the same direction as B_E or opposite it. The same total magnetic field, B_T, is required for both tube orientations to maintain the same orbit radius. However, the required coil current is different in the two cases. Let the larger current be I_ℓ and the smaller be I_s. In the first case the Earth's magnetic field B_E is opposing the field generated by the coil and in the second case the Earth's field is adding to the field generated by the coil. Considering both cases in turn, we have for the total field B_T

$$B_T = B_l - B_E = K_r I_l - B_E \qquad (7)$$

and

$$B_T = B_s + B_E = K_r I_s + B_E \qquad (8)$$

These can be added to give

$$B_T = \frac{K_r}{2}(I_l + I_s) \qquad (9)$$

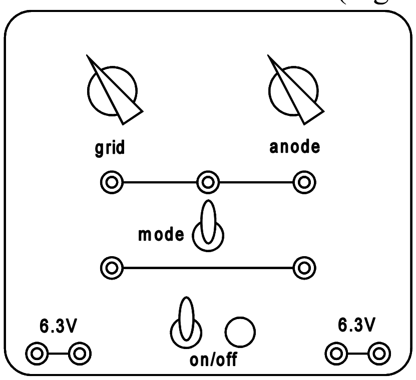
and subtracted to give

$$B_E = \frac{K_r}{2}(I_l - I_s) \qquad (10)$$

B_T will be used in equation (3) to determine *e/m*.

1. Setup

a. Locate the power supply manufactured by Leybold (see Fig. 4a). The mode switch is to remain in the up position; it is a range selector and we will be using the upper 0-25 V and 0-300 V ranges. The anode voltage controls both the beam intensity and the diameter of the electron orbit. For optimum beam sharpness the anode voltage should be between 150 and 250V. Beware, to increase the anode voltage you rotate the knob *clockwise*; to increase the grid bias you rotate the knob *counter-clockwise.* Before switching on, turn both knobs to zero. After a few minutes, the filament will warm up and emit a yellow glow. Once it is hot, increase the anode voltage until the beam becomes visible (Fig 4b). Change grid bias slightly to obtain best beam definition.



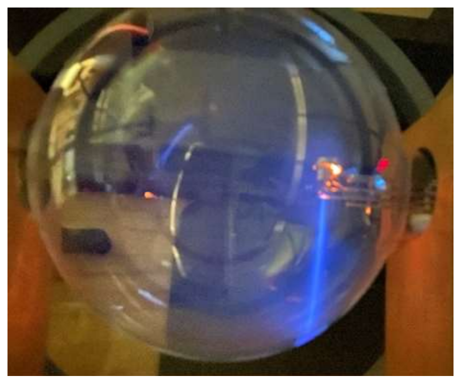


Fig.4 a) Schematic of the Leybold power supply b) beam with no magnetic field

b. The DC current for the Helmholtz field coils is adjusted by second power supply, with a rheostat in series (used as a protection resistor in this case) and a knife switch that allows the direction of current to be reversed, when the knife switch is vertical there is an open circuit. By increasing the current in the coils you should be able to see the electron beam bend. Rotate the tube until the beam there is no component of the beam parallel to the magnetic field causing to spiral as seen in Figure 5.

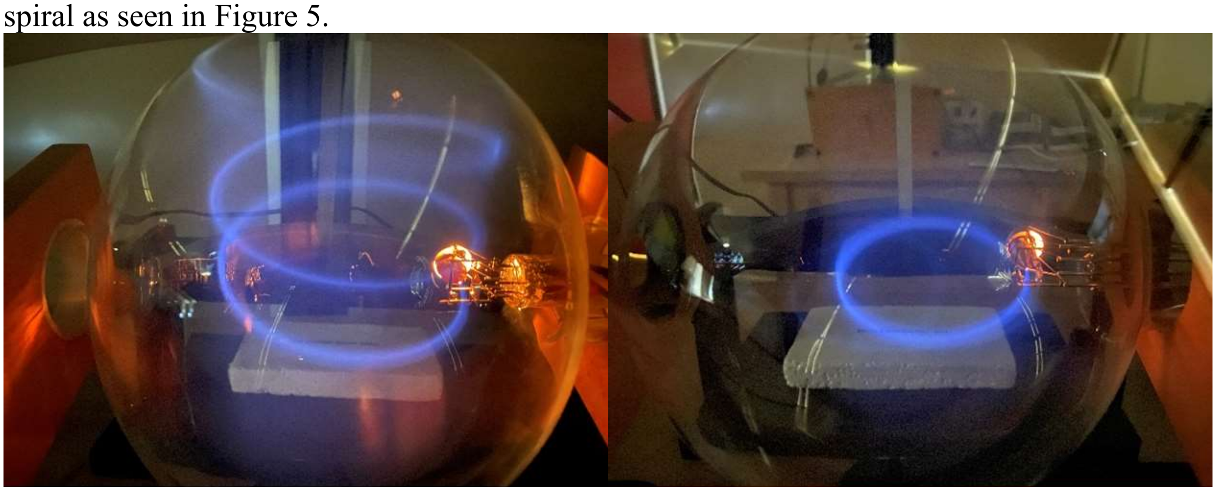


Fig. 5 Beam with component of velocity parallel to field, and rotated so beam is perpendicular to field

c. The diameter of the beam circle is measured using the illuminated scale. Adjust the position of the scale so that the images formed by the two surfaces of the plate glass appear parallel to and on each side of the plane of the electron beam, and along a diameter of the circular path of the beam. This is best determined by viewing from the side. If the scale is not in the plane of the beam then there will be a systematic error in the diameter measurements. Figure 6 shows the scale set up at two distances above the apparatus producing different measured diameters for the same actual diameter. Once the position of the scale is adjusted all measurements are made by looking down and lining up the two images of the scale to avoid parallax as shown in Figure 7.

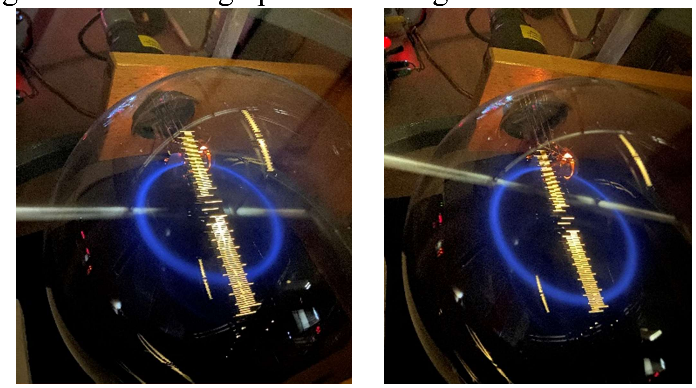


Fig. 6 Scale set up at two distances above the apparatus. The actual beam diameter is the same in both cases

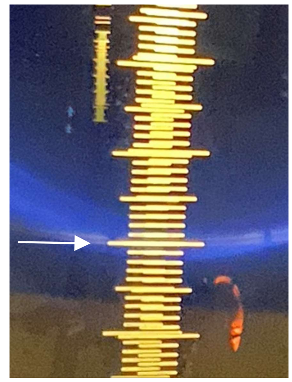


Fig 7. View of scale from top with two images of light scale lined up

2. Accelerating voltages and coil currents for 8cm and 10cm diameter orbits

1. Set the anode voltage (V) to 150 V and increase the current in the Helmholtz coil until the diameter of the electron orbit (D) is 8 cm. Errors in the digital meters should be taken as ±1 in the last digit. Rotate the tube gently in its wooden clamps about its long axis so that the beam forms a closed circle rather than a spiral. The spiral is of interest because it demonstrates that a velocity component parallel to the magnetic field is unaffected.
2. Increase the anode voltage to 250 V in 20 V increments, measuring the current required to maintain the electrons in a circular orbit with D = 8 cm. How accurately can the diameter be reproduced and measured?
3. Turn the tube through 180 degrees and also reverse the direction of the current in the coils using the reversing switch. This bends the beam into a circle with the direction of electron motion reversed.
4. Repeat steps 1-2 with the polarity of the current reversed.
5. Repeat 1-4 for $D = 10$ cm. which, if either, of the two orbit diameters is closest to being symmetric about the centre of the coils? Is this an important or negligible effect?

After the experiment, do not disconnect the leads but ensure that the power supply and the coil current are switched off. The reversing switch should be in the central position.

Before you leave use one set of measurements (matching I_l and I_s values) to estimate *e/m* and B_E. If your values differ from accepted values by more than 10% you should check your measurements. *Also keep in mind that you measure diameter but calculations are in terms of r.*

Analysis and Report

Subtract 1% from your anode voltage measurements to correct for the voltage drop in a protection resistor included in the leads.

For each beam diameter, determine *Kr* and make the appropriate additions and subtractions of the coil currents at each anode voltage to obtain B_T and B_E using Eqs. (9) and (10).

For each measurement pair, use B_T to calculate e/m. Average your resulting values to obtain the best estimate of e/m. Compare your experimental determination of e/m with the accepted value.

Calculate the Earth's magnetic field by averaging all estimates of B_E. The apparatus has been oriented along the Earth's field so your measurement should give the magnitude of the field. If possible, compare the Earth's magnetic field with that expected value at this latitude and longitude. (Links to Geophysics sites can be found on the course webpage).

Derive Eq. (4) and use the dimensions given to calculate K. Does this agree with the value given by the manufacturer?

Recall that the uncertainty in an average is the standard error i.e. standard deviation of the mean (not the average of the propagated errors).

Ferromagnetic Hysteresis

Hysteresis is the dependence of a system not only on its current environment but also on its past environment. The magnetic hysteresis curves of two different ferromagnetic materials, transformer iron and carbon steel, are investigated.

Introduction

Ferromagnetic materials are indispensable for electrical machines, transformers and many other devices. These materials have the remarkable characteristic that they increase the magnetic field **B** in magnetic circuits by many orders of magnitude compared to the same circuits constructed with non-ferromagnetic cores. In ferromagnetic materials, due to a quantum mechanical phenomenon, the unpaired electron spins line up parallel with each other over small volumes called domains. The electron spins act like current loops to produce a magnetic field and as a result each domain acts like a small magnet. Domains, which range from 0.1 mm to a few mm in size, are generally not aligned with each other so a sample of ferromagnetic material generally produces no net magnetic field. An externally imposed magnetic field $\mathbf{B_0}$, from a solenoid for example, can cause the magnetic domains to line up with each other and the material is said to be magnetized. The internal magnetic fields which come from the alignment of the electron spins can be hundreds of times stronger than the external magnetic field required to produce the changes in domain alignment. The effective multiplication of the external field $\mathbf{B_0}$ which can be achieved by the alignment of the domains to produce the net field **B** is often expressed in terms of the relative permeability μ_r[1]

$$\mathbf{B} = \mu_r \mathbf{B_0} \quad (1)$$

The external field used to align the domains can be written as

$$\mathbf{B_0} = \mu_0 \boldsymbol{H} \quad (2)$$

so

$$\mathbf{B} = \mu_o \mu_r \mathbf{H} \quad (3)$$

where μ_0 is the relative permeability of free space and $\mathbf{H}$, often called the magnetizing force, is related to the magnetizing current. In some simple cases, H can be obtained from Ampere's law

$$\oint H \cdot d1 = NI \quad (4)$$

where N is the number of turns (coils) around the path and I the current in each turn. For the case of the toroid shown in Figure 1, $LH = NI$ where L is measured around one of the flux lines, so $B = \mu_o \mu_r NI/L$. This is an approximation since

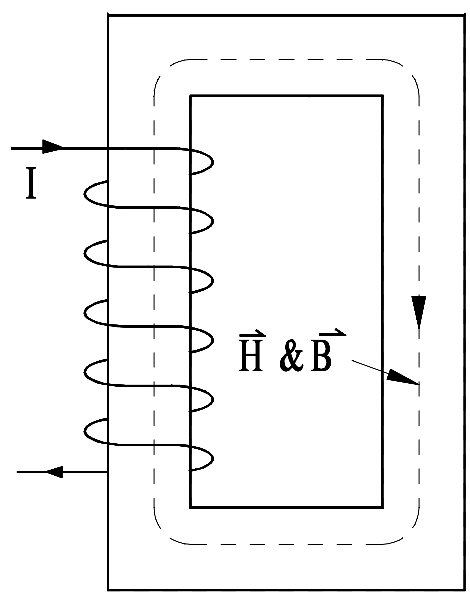


Figure 1. Magnetized Toroid

[1] Note that μ_r is not a constant but varies with the magnetizing current.

L is well defined only for a toroid of small cross sectional area. For this case $\mu_o H$ is just the flux density one would obtain with a vacuum for the core of the toroid (assuming a large number of coils uniformly distributed around the toroid) so that μ_r is the relative increase in flux density caused by the presence of the material of the core. For a vacuum $\mu_r = 1$, and μ_r is essentially 1 for all non-ferromagnetics, e.g., Al, air, wood, plastic etc., but for ferromagnetics μ_r can range up to nearly 10^6, with values of 10^3 being common.

Along with this enormous increase in magnetic field, a complication occurs. It is found that when I is varied, B depends not only on the instantaneous value of H (or I) but also on the magnetic history of the sample. (From now on only the magnitudes of **B** and **H** will be considered). If we start with an unmagnetized ferromagnetic core and apply an increasing H, by increasing I, the induced B will follow the initial magnetization curve shown in Figure 2. This curve levels out when all the domains are aligned – a situation referred to as saturation. If the current in the magnetizing coil is then reduced to zero some domains remain aligned and the field at this point B_r is called the remanence. To reduce the magnetization to zero the current must be reversed to produce a negative H value H_c the coercive force. At this point enough of the domains have been reversed to cancel the effect of those remaining aligned. If the current is periodically reversed B will follow the hysteresis loop shown in Figure 2. A representation of the corresponding sequence of domain alignments is shown in Figure 3.

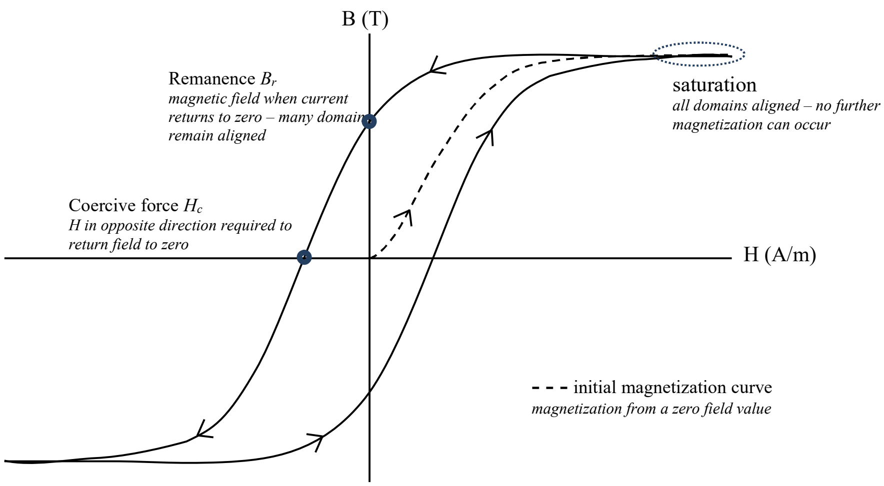


Figure 2. Hysteresis Curve

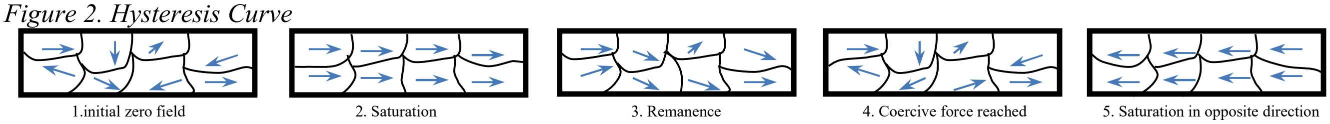


Figure 3: Sequence of domain alignments

This cyclic behaviour is called hysteresis and results from a kind of internal friction as the domains change. Similar behaviour is encountered in other fields e.g. in ferroelectricity, and elasticity. As the domains are changed energy is lost to heat. It can be shown that the heat produced with each complete cycle, in J m^{-3}, equals the area of the loop (when B is in T and H in Am^{-1}). The power dissipation in the sample is therefore given by the product of the area of the hysteresis loop A, the volume of the sample being magnetized V, and the frequency f with which the loop is cycled, i.e.,

$$P = AVf. \quad (5)$$

The shape of the hysteresis curve and the size of B_r and H_c determine a material's suitability for an application. For example, to make a permanent magnet or for use in magnetic memory (hard drive or magnetic strip on a credit card) a large H_c is desirable so that the material is not easily demagnetized. For generators, motors and transformers the hysteresis curve should be tall and thin indicating that a) large B values are obtained with small magnetizing currents, b) H_c values are small so that the field is easily reversed and c) the loop area is small indicating low energy loss.

As mentioned previously μ_r varies with the magnetizing current. $\mu_r(H)$ can be determined from the initial magnetization curve by rearranging equation (3) as $\mu_r = \frac{B}{H\mu_0}$. The value of μ_r will reach a maximum as saturation starts to be approached.

More information can be found in *Applied Electromagnetics* by M. A. Plonus, pp.350-357. (see moodle)

Experiment

An oscilloscope is used to examine hysteresis loops for two samples, iron and steel. The circuit diagram of the apparatus is shown in Figure 4. The material under test is made a part of the ferromagnetic core of a transformer. The primary (i.e. input) winding has n_1 (160) turns and carries an oscillatory current I.

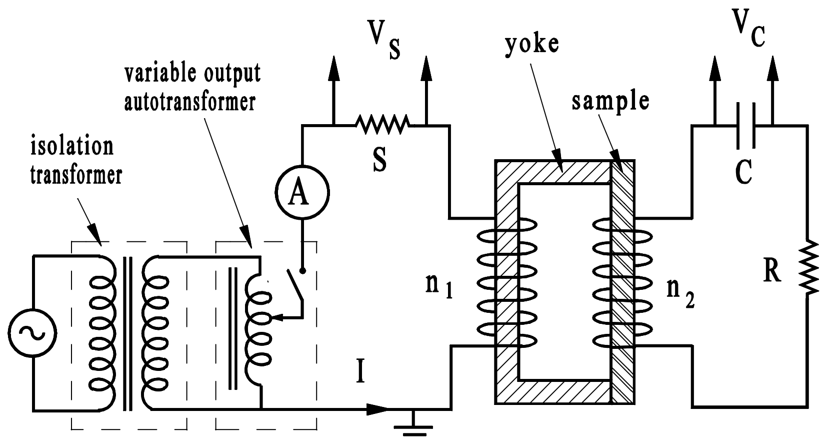


Figure 4. Experimental Circuit

The series resistance ($S = 0.10\,\Omega \pm 5\%$) provides a voltage signal V_s proportional to the instantaneous value of I. If the material of the core is the same everywhere, then H can be calculated from equation (4) as $H = n_1 I/L$ where L is the magnetic length of the core. (The case of two different materials in the core is dealt with through the determination of L as described in the procedure). Therefore,

$$H = \frac{n_1}{LS} V_s. \quad (6)$$

V_s is used to drive the X axis of the oscilloscope. The ammeter is used only as an approximate indicator and not as a measuring device. Like most AC meters, it reads RMS values. However, the current is not sinusoidal and so the calibration factor is not known.
The secondary coil, with n_2 (150) turns, has induced EMF e due to the flux changes within the sample, i.e., $e = d\phi/dt = n_2 A_c (dB/dt)$ where A_c is the cross sectional area of the sample and B the mean flux density within it. The RC ($R = 1.00\,M\Omega \pm 1\%$, $C = 0.50\,\mu F \pm 2\%$) circuit integrates e and the output of V_c is given by

$$V_c = \frac{q}{C} = \frac{1}{C}\int i dt \approx \frac{1}{RC}\int e dt = \frac{1}{RC}\int d\phi = \frac{n_2 A_c B}{RC}. \quad (7)$$

This equation uses the approximation that the current $I \approx e/R$ which is valid provided $V_c << e$. This condition is satisfied in this experiment. Therefore,

$$B = \frac{RC}{n_2 A_c} V_c. \quad (8)$$

V_c is used to drive the Y axis of the oscilloscope.

Procedure

A: Investigate shape of hysteresis curve

1. Adjust the trigger on the oscilloscope → *Trigger Menu → Type = edge, Source = AC Line (signal from wall outlet), Note: if signal jumps around reset these.*
Zero both channels by depressing the Vertical Position buttons and check that both channels are set for x1 probe. Set the oscilloscope to display Channel 2 as a function of Channel 1 (XY Mode) by selecting: *Horizontal Menu → XY*.

2. Connect the leads to the scope such that H is displayed on the X-axis and the B-field is displayed on the Y-axis. The horizontal and vertical sensitivity should be set to 200mV per division as a starting point.

3. Place the transformer iron sample on the yoke, set the variable autotransformer (variac) control to zero and switch it on. Gradually increase the voltage on the variac. A hysteresis loop like the one in Figure 2 should begin to develop on the screen. The signal will likely be very noisy, but this can be compensated for by having the oscilloscope average the signal over many samples. The averaging function is found in *Waveform Acquire→Mode =Average, set number of cycles to average over*.

4. Increase the voltage until the loop spans four divisions (± 2) of the horizontal axis or 400mV (if you have altered the voltage base). This will correspond to a magnetization current of 4 Amps peak. Make note of the hysteresis being displayed as well as the variac level.

5. Switch the oscilloscope back to time domain (*Horizontal → Main/Delayed → Main*). Adjust the time resolution so that the display shows 5ms per division.

6. Collect a *Comma-Separated Values* (CSV) file of the waveform by connecting a USB stick then: → *Save/Load → External storage→ New → New File → Save as = CSV → Save.* Go to a computer in the lab and open the file in Excel to check your data. If it looks the same as on the oscilloscope email it to yourself.

B: Investigate initial magnetization curve

7. Return the display to XY mode. Starting from zero, gradually increase the variac voltage while recording the progression of the upper right corner of the the hysteresis loop. Fairly accurate measurements can be made using the oscilloscope cursors: → *Cursor → Source = CH1 → Adjust to move cursor, tap Adjust button to switch between cursor 1, 2 and both together*. Do the same for *CH2.* To increase accuracy while measuring the corner of the loop, it is recommended that you adjust the scales of the oscilloscope to provide a clearer picture. Be sure to take an adequate number of points at the lower levels so that an accurate trace of the initial magnetization curve is obtained. The full data set should consist of at least 15 points.

Turn off the variac without reducing the voltage to zero. Gently try lifting the iron sample.
What happens? Why? Turn the variac back on and reduce the voltage to zero. Replace the iron sample with the carbon steel sample.

8. Repeat steps 3 to 6 for the carbon steel sample (do not repeat step 7).

C: Qualitative investigation of power loss

9. When you are done recording data for the carbon steel sample touch the sample and observe that it has heated up due to energy losses.
   a. Record a rough estimate of what wattage bulb you feel would produce that much heat so that you have a point of comparison for your energy loss calculations. If you have never touched a lit bulb try touching the one by your lab station.
   b. Touch the middle of the steel sample, the ends of the steel sample and the iron yoke. Record your observations. Based on your observations what length should be used to determine the volume to use for energy loss calculations (see Figure 5).

D: Apparatus measurements

10. Take the following measurements of the apparatus[2]:
- A_c : The cross-sectional area of the samples. This is simply the cross-section of the bars placed on the yoke.
- L : The magnetic length of the samples.

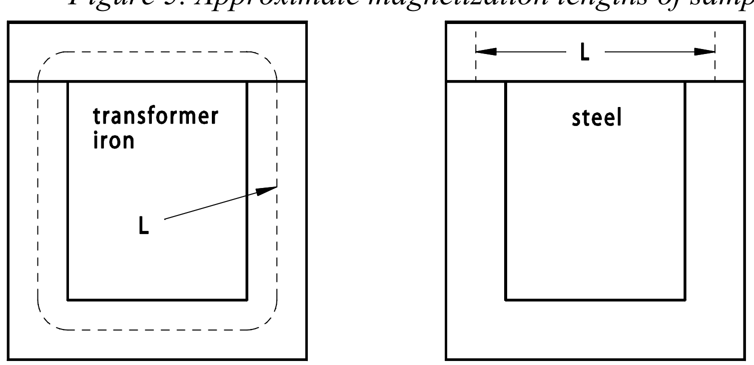
Transformer Iron:

The yoke and the sample are made of the same material, so use the combined length of the sample and yoke for length L. Notice from Figure 5 that this length is somewhere between the inner and outer perimeters. Use a path as outlined in the sketch. Although physical lengths are measured quite accurately, estimating the magnetic lengths L and area A is clearly more difficult. Try to make sensible estimates for the uncertainties in these quantities. If this creates a problem, discuss it with a lab demonstrator.

Carbon Steel:

In this case the yoke is much more easily magnetized than the sample. This means that the value of H in the sample is much greater than in the yoke (B being the same in both) and as a simplifying approximation we shall assume that $H = NI/L$ where L is now the magnetized length of the sample only. Use the length shown in Figure 5 and again try to estimate reasonable uncertainties for L and A_c.

Figure 5. Approximate magnetization lengths of samples



[2] An explanation of how to read the Vernier Calipers can be found in the Ultrasound experiment. Note that measurement of height can be made using the thin rod that sticks out the end.

E: Investigate effect of air gap and eddy currents

11. Experiment with placing the plastic and copper spacers between the transformer iron and the yoke. Introducing these spacers will cause a change in the shape of the hysteresis loop. The plastic spacer will act as an air gap, causing the same magnetizing current to produce much less hysteresis. In general, the larger the air gap, the better the linearity. The copper spacer is available to show the effects of eddy currents (used in magnetic brakes). Save traces for each type of spacer so you will be able to discuss them.

Analysis and Report

1. Using the data saved in step 6 determine B and H from equations 6 and 8 and plot the hysteresis loops for both the transformer iron and carbon steel on the same graph. With the plots, determine the remanence and coercive force for both samples.

2. Of the two materials investigated, which would be most effective, and why, for the following applications:
   a. A permanent magnet or a magnetic memory device.
   b. A motor.

3. Determine the power loss in the carbon steel hysteresis loop using the data from step 6 and Equation 5. Use the magnetizing length, L, to calculate the volume in which the loss occurred.
NOTE: Determining the area of the hysteresis loop is most easily accomplished by taking data that encompasses <u>one</u> hysteresis loop and using numerical integration techniques (Simpson's Rule or Riemann Sums) to determine the area enclosed.
   a. Compare the value you obtain to your estimation of power dissipation in step 9 a).
   b. Based on your observations of the steel sample and yoke in 9b) did the use of the magnetizing length in the calculation of the volume for the hysteresis loss in the steel appear to be justified?

4. For Iron, plot B versus the H for the initial magnetization curve using the data from step 7. Using Equation 3, plot a curve of relative permeability (μ_r) versus the H field.
   a. What is the maximum relative permeability? Look up the accepted value and compare.
   b. At what flux density does it occur?

7. Discuss the hysteresis curves with the plastic and copper spacers. Consider the effect on energy loss and maximum magnetic field.

Damped Harmonic Motion

This is an investigation of damped harmonic motion in both a mechanical and an electrical system under various damping conditions.

Introduction

Oscillating systems almost invariably have damping which serves to reduce the oscillation amplitude with time. Such systems are found throughout physics and engineering. Atomic force microscopes can view surfaces at the atomic level and typically contain a small cantilever which is a mechanical oscillator with damping. Automobiles suspensions are damped so that oscillations induced by a pothole are minimized. In this case the damping is so strong that the oscillations may no longer be visible. This experiment introduces two examples, a mechanical system and an electrical system, both of which are described by the same equations. In each case we examine the effects of varying the strength of damping over a wide range.

Theory

In the absence of a driving force, the damped oscillator equation is conventionally written in the following form

$$\ddot{x} + \gamma\dot{x} + \omega_o^2 x = 0 \quad (1)$$

where we assume the damping is proportional to the velocity $\dot{x}$, the strength being given by the constant γ. In the absence of damping, i.e. when $\gamma = 0$, the equation for simple harmonic motion (SHM) is retrieved

$$\ddot{x} + \omega_o^2 x = 0 \quad (2)$$

where ω_o is the angular frequency of oscillation. As usual, $\omega_o = 2\pi f_o = 2\pi/T_o$.

The best chance of finding an analytical solution to Eq. 1 is to find a function which, upon repeated differentiation, essentially reproduces its own form which can then be cancelled out as a factor. For example we could try the exponential function

$$x(t) = Ae^{\lambda t}, \quad (3)$$

where λ is still to be determined. Substituting this trial solution into Eq. 1 leads to

$$(\lambda^2 + \gamma\lambda + \omega_o^2)Ae^{\lambda t} = 0. \quad (4)$$

Consequently, the exponential solution will be acceptable provided that the auxiliary equation

$$\lambda^2 + \gamma\lambda + \omega_o^2 = 0 \quad (5)$$

(which is an ordinary quadratic in λ) is satisfied. In general there are 2 values λ_1 and λ_2 which do this

$$\lambda_{1,2} = -\gamma/2 \pm \sqrt{(\gamma/2)^2 - \omega_o^2}. \quad (6)$$

Because the damped oscillator equation is linear, any linear combination of the above two solutions is also a solution, i.e.

$$x(t) = Ae^{\lambda_1 t} + Be^{\lambda_2 t} \quad (7)$$

and this is the most general solution of Eq. 1. It contains 2 arbitrary constants A and B (which are fixed by the initial conditions) and remains valid even if $\lambda_{1,2}$ are complex, the only exception being when $\lambda_1 = \lambda_2$, i.e. $\gamma/2 = \omega_o$ in Eqn. 6. In this special case, the general solution is

$$x(t) = (A + Bt)e^{-\frac{\gamma}{2}t}. \quad (8)$$

which you can prove by direct substitution in Eq. 1.

There are three different possibilities for the form of the solution and we will now examine these in more detail. Figs. 1 and 2 illustrate the results. Fig. 1 is appropriate to a system in equilibrium at $x = 0$ which is given an impulse (a velocity $\dot{x} = 50$ units/sec) at $t = 0$, i.e. similar to the mechanical system in this lab. Fig. 2 is for a system which is in equilibrium at $x = 1$ unit, and then changed so that the new equilibrium value is $x = 0$, i.e. similar to the electrical case used in this lab.

(i) Underdamped Solution ($\gamma/2 < \omega_o$)

In this case $\sqrt{(\gamma/2)^2 - \omega_o^2}$ is imaginary and we write

$$\sqrt{(\gamma/2)^2 - \omega_o^2} = j\sqrt{\omega_o^2 - (\gamma/2)^2} = j\omega \qquad (9)$$

where $j = \sqrt{-1}$. It follows from this definition that, in the absence of damping, $\omega = \omega_o$. The solution is of the form

$$x(t) = Ae^{(-\frac{\gamma}{2}+j\omega)t} + Be^{(-\frac{\gamma}{2}-j\omega)t} \qquad (10)$$

$$\text{or} \qquad x(t) = e^{-\frac{\gamma}{2}t}\{C\cos\omega t + D\sin\omega t\}. \qquad (11)$$

Thus $x(t)$ is oscillatory, with angular frequency ω and an exponentially decaying amplitude. The motion is illustrated in Fig. 1a and 2a. (As an example of the calculation, if $x = 0$ and $\dot{x} = v_o$ at $t = 0$ then $C = 0$ and $D = v_o / \omega$. Thus $x = (v_o / \omega)e^{-\frac{\gamma}{2}t} \sin \omega t$ which is plotted in Fig. 1a). The maxima and minima occur when $\dot{x} = 0$, which leads to $\tan\omega t = 2\omega/\gamma$ = constant, i.e. adjacent maxima and minima are separated by $T/2 = \pi/\omega$. The positions of successive maxima and minima and the times at which they occur are indicated in Fig. 1a and 2a by x_1, x_2, x_3 … and t_1, t_2, t_3 ….

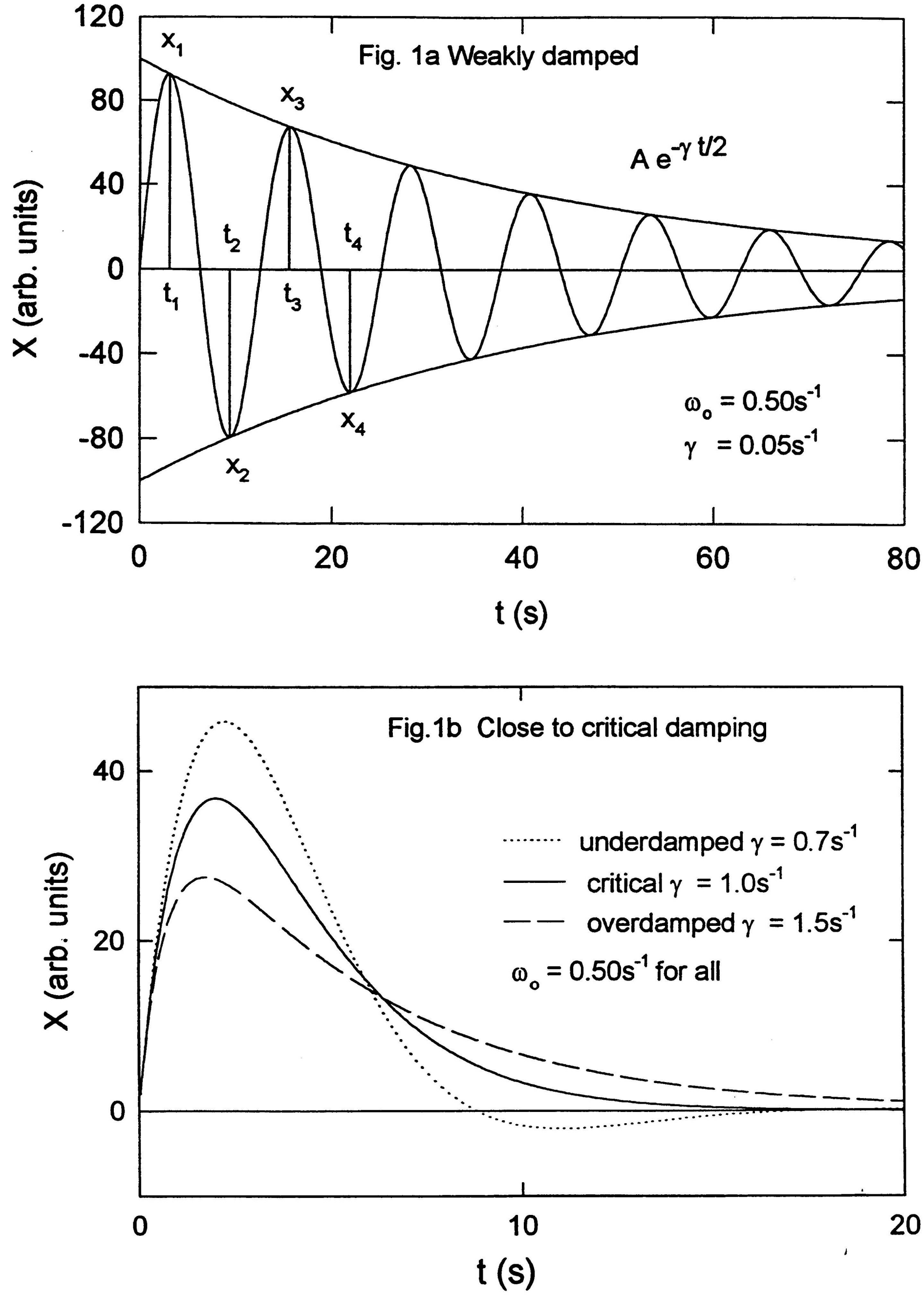


Fig. 1 These plots represent the behaviour of the mechanical system.
In these calculated curves the system was started at x=0 with a velocity of 50 units/sec. Notice the scale changes in the lower panel.

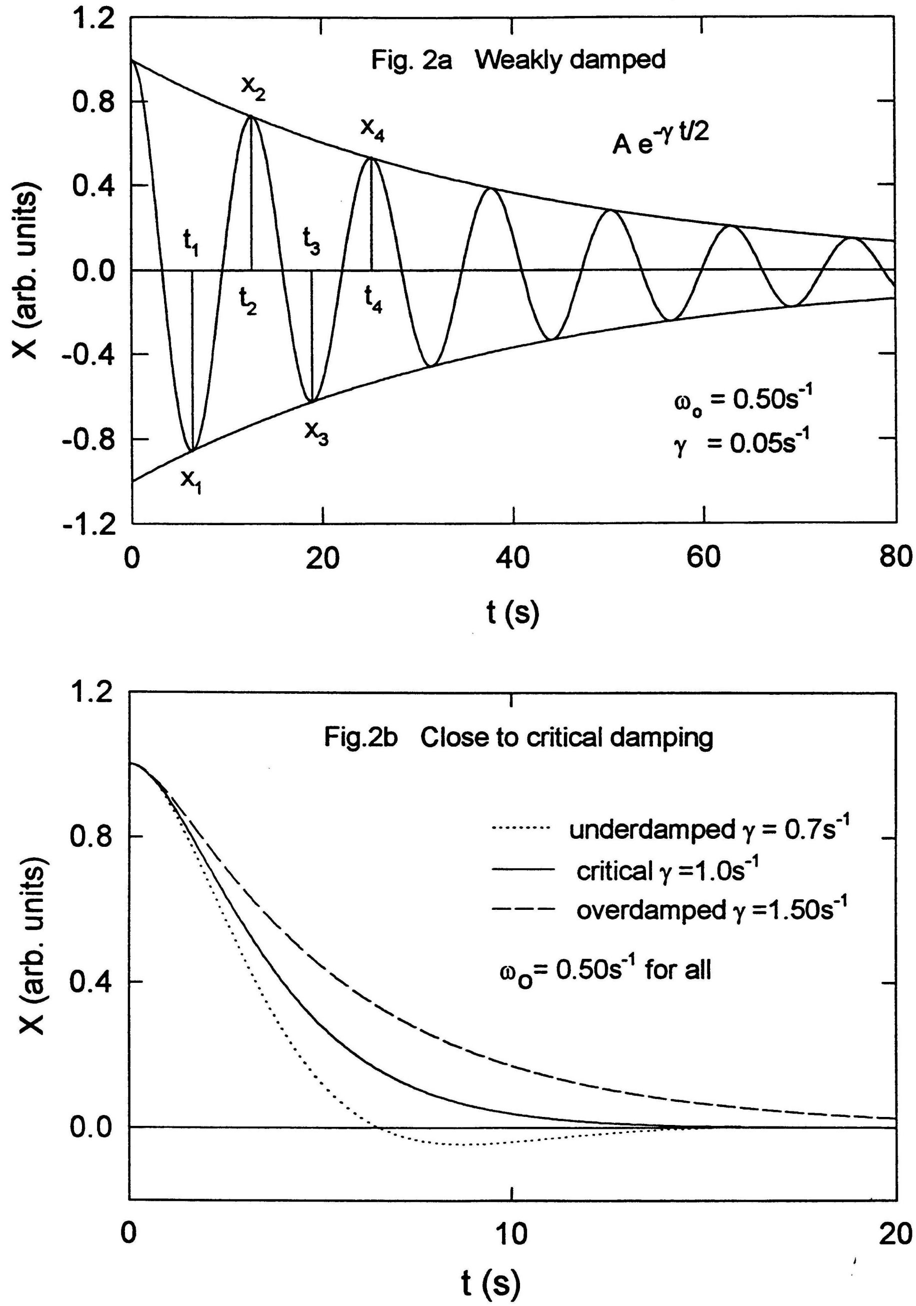


Fig. 2. These plots represent the behaviour of the electrical system.
In these calculated curves the system was started with a displacement of 1 unit and zero velocity. Notice the horizontal scale change in the lower panel.

(ii) Overdamped Solution ($\gamma/2 > \omega_o$)

Now $\sqrt{(\gamma/2)^2 - \omega_o^2}$ is real, say β (which is equivalent to $j\omega$), so the solution is now the sum of two exponentials with real exponents

$$x(t) = Ae^{(-\frac{\gamma}{2}+\beta)t} + Be^{(-\frac{\gamma}{2}-\beta)t}. \quad (12)$$

The motion is not oscillatory and is illustrated in Fig. 1b and 2b by the dashed lines (e.g. the overdamped curve in Fig. 1b is $x = (v_o / \beta)e^{-\frac{\gamma}{2}t}$ sinhβt).

(iii) Critical Damping $(\gamma/2 = \omega_o)$

The solution is given in Eq. 8. The motion is non-oscillatory and the system returns to equilibrium in the minimum time (but this is not proved here). The full curves in Fig. 1b and 2b illustrate critical damping (e.g. the critical curve in Fig. 1b is $x = v_o te^{-\frac{\gamma}{2}t}$). It is desirable in many mechanical and electrical systems to have critical damping since this allows them to adapt to a new equilibrium or to recover from an impulse in the minimum time.

Experiments

1. Mechanical System

The torsional oscillations of a ballistic galvanometer are studied first. A ballistic galvanometer is essentially a coil which is suspended by a fibre in a radial magnetic field. A small mirror is attached to the coil to enable the angular position θ to be measured by viewing the reflection of a scale mounted around the apparatus. Examine the instrument provided and sketch the key features.

An angular impulse on the coil is created due to the magnetic field when a current is sent through the coil. Subsequently, the rotation of the coil in the magnetic field induces an emf $\mathcal{E}$ which produces a current in the coil if an external resistance R is connected across it. This dissipates energy at a rate $\mathcal{E}^2$/R and gives damping. Notice the damping increases as R is decreased. After being disturbed, the subsequent variation of θ with time can be well-described by the damped oscillator equation, Eq. 1. (See Bleaney and Bleaney Electricity and Magnetism).

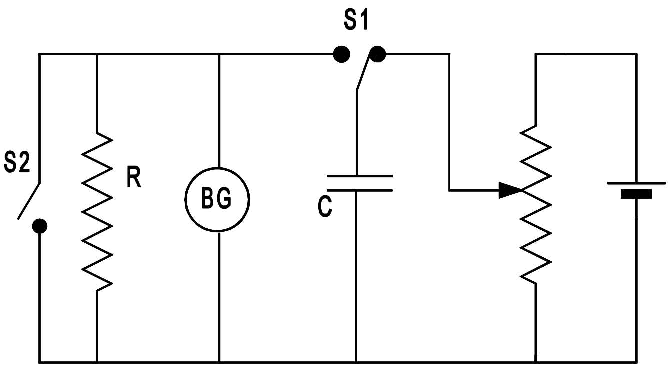


Fig. 3

The galvanometer (BG) is connected in the circuit shown in Fig. 3. The capacitor is charged to a fraction of the voltage of the cell determined by the setting of the potentiometer. When switch $S1$ is moved to the other position, the capacitor is discharged producing a current through the galvanometer coil giving it an initial angular impulse. The potentiometer should be adjusted such that the initial displacement goes to the end of the scale but not past. The parallel resistance R supplies the variable damping. When required, the switch $S2$ short-circuits the galvanometer coil, providing maximum damping and allowing the coil motion to be arrested quickly. In this experiment the angle θ of the galvanometer coil is observed by viewing the reflection of the scale, which arches around the experiment, in the mirror attached to the coil. The scale reading x can be taken to be accurately proportional to θ.

By changing the magnitude of R, examine the three different solutions to the damped oscillator equation as in the following. (Note that Figs. 1(a) and 1(b) are appropriate to this system).

- Underdamped

(i) Use a large value of R, even $R = \infty$ is satisfactory though in this case viscous damping due to the air will be the relevant mechanism. Measure the period T so you can determine $\omega = 2\pi/T$. It is most accurate to measure the time for several periods and divide by the number of cycles.
(ii) To determine γ proceed as follows. Measure the amplitude x_n of the oscillation at successive maxima and minima, numbering the maxima and minima from 1 to n respectively (cf Fig. 1(a)). According to Eq. 11, the amplitude of successive maxima and minima should decay exponentially with a time constant of $2/\gamma$. A plot of $\ln x_n$ vs t has a slope of $\frac{\gamma}{2}$. Having obtained ω and γ, ω_0 can be calculated from equation (9).

- Critical Damping

Adjust the resistance R so that the system is critically damped. There should be no overshoot, but a small increase in R should produce overshoot (cf Fig. 1(b)). Measure R and the time from the impulse until the deflection is a maximum. Using Eqn. 8 and noting $x = 0$ and $\dot{x} = v_o$ at $t = 0$, one finds $x = v_o t e^{-\frac{\gamma}{2}t}$. Using $\dot{x} = 0$ at maximum deflection, γ can be determined from this equation using the time to maximum deflection and compared with the prediction $\gamma = 2\omega_o$.

- Overdamped

Theory predicts that as damping is increased the initial displacement is less but the coil takes longer and longer to return to its initial position as shown in Fig. 1(b). Reduce R to about a third of the value found for critical damping. Qualitatively examine and comment on the displacement pattern (make notes in your lab book).

2. Electrical System

The circuit is shown in Fig. 4 and consists of a signal generator of output resistance R_G, a capacitor C, an inductor L with resistance R_L, and a variable resistor R_v connected in series. The generator produces a square wave which periodically switches the voltage applied to the circuit between two steady values, say $V_{1,2}$. The voltage across the capacitor is measured with the oscilloscope. Note that the generator has the capability of supplying a dc offset voltage to the signal. We suggest that you use the dc input option on the oscilloscope and centre the trace using the dc offset on the generator. The ac option on the oscilloscope may lead to 'droop' of the signal as a function of time. Ask if in doubt.

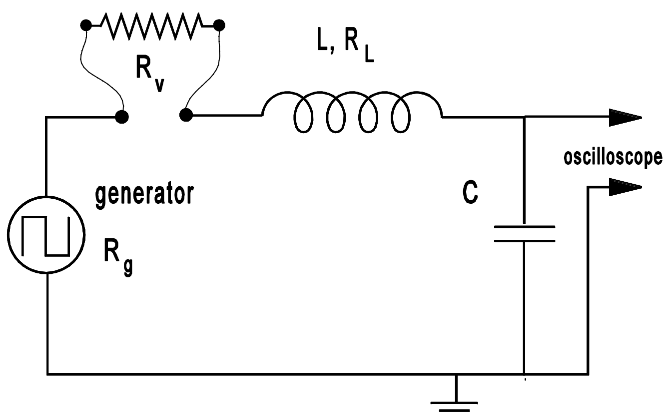


Fig. 4

For the moment, assume that only a constant voltage V_1 is applied to the circuit by the generator so that the sum of the voltages around the circuit is V_1. Expressing this in terms of the charge q on the capacitor, and noting that the current $I = \dot{q}$ we have

$$L\ddot{q} + R\dot{q} + q/C = V_1 \quad (13)$$

where $R = R_L + R_G + R_v$. Using a new variable $Q = q - CV_1$, the resulting differential equation for Q is of exactly the same form as Eqn. 1, and we see

$$\gamma = R/L \quad (14)$$

and

$$\omega_o^2 = 1/LC \quad (15).$$

The solution for q, and hence the voltage qC on the capacitor, is of the same form as Eqn. 7, except an extra term CV_1 is added to account for the steady voltage V_1. Each time the generator voltage switches between $V_{1,2}$, the capacitor voltage will show the 'transient' solution of Eqn. 7 superposed on the next value of $V_{1,2}$. The period of the square wave is made long enough for the transient to die away before the next voltage step. Figs. 2(a) and 2(b) are appropriate to this system.

- Underdamped

(i) Set the variable resistor R_v to its minimum value (which corresponds to minimum damping), and obtain a stationary (oscillatory) waveform by adjusting the trigger controls on the oscilloscope. → *Trigger Menu → Type = edge, Source = CH1 (no external trigger on this scope) → Trigger Level, adjust this knob until the signal is stable – the level is shown as a line on the screen and should be within the range of the signal i.e. not above or below it or too close to the top or bottom.*
Measure the period T so you can determine $\omega = 2\pi/T$. It is most accurate to measure the time for several periods and divide by the number of cycles.
→ *Cursors → Time → Multipurpose to move cursor, tap Multipurpose to switch between cursors. Note the fine button allows you to make fine adjustments to the cursor position.*

(ii) To determine γ proceed as follows. Measure the amplitude x_n of the oscillation at successive maxima and minima, numbering the maxima and minima from 1 to n respectively (cf Fig. 2(a)): → *Cursors → Amplitude, Set one cursor to steady state level and move other cursor, the ΔH will give you the distance above or below steady state.*
According to Eq. 11, the amplitude of successive maxima and minima should decay exponentially with a time constant of $2/\gamma$. A plot of $\ln x_n$ vs t has a slope of $\frac{\gamma}{2}$. Having obtained ω and γ, ω_0 can be calculated from equation (9).
(iii) To calculate the theoretical value expected for ω_o from equation (15) you will need to record the values for L and C (given on the apparatus). To determine the expected value of γ from equation (14), you will also need the total resistance R. Disconnect the circuit from the signal generator and oscilloscope and use a multimeter to measure R_v (which should be approximately zero) and R_L. The total resistance in the circuit also includes the internal resistance of the signal generator R_g. The following explains how to measure this. Disconnect the generator from the circuit and connect it directly to the input of the oscilloscope. Adjust the gain until the step in the square wave is approximately full scale. Without altering the output of the generator, connect a resistor box across the generator and adjust the resistance until the vertical deflection on the oscilloscope is exactly half its original size. The internal resistance of the generator R_g is equal to the setting of the resistor box. You should prove this in an Appendix.

- Critical Damping

Reconnect the signal generator and oscilloscope. Increase R_v and observe the behaviour of the waveform as R_v is increased (comparing with Fig. 2b). Find the smallest value of R_v for which there is no overshoot (cf Fig. 2(b)). You will compare the measured value of the total resistance with the theoretical value. This latter is obtained from the condition $\gamma = 2\omega_o$ which is equivalent to $R^2 = 4L/C$. In practice it is difficult to distinguish critical damping from underdamped cases close to critical, e.g., the amplitude of the peak at $t_1 = T/2$ in Fig. 2 for $\gamma = 0.8 \times 2\omega_o$ is only 1% of the initial displacement at $t = 0$. For this reason it is important to get a good estimate of uncertainty on R_v and thus for R.

- Overdamped

As R_v is further increased, qualitatively examine the waveform (using Fig. 2b as a guide). Theory predicts that the output voltage takes longer and longer to return to the steady state value as in Fig. 2(b).

Data capture so graphs can be compared:
Go back to the R_v values used in each case above and save the data so that you can produce comparison graphs similar to those in Figure 2.
Collect a *Comma-Separated Values* (CSV) file on USB stick of each waveform by

1. *→ Resources Save/Recall → Action = Save Waveform (rotate Multipurpose to move to this, press Multipurpose to select) , Save to = USB File (*.csv) → Save*
2. Go to a computer in the lab and open the file in Excel to check your data. If it looks the same as on the oscilloscope email it to yourself.

Analysis

Mechanical system

- Underdamped

(i) Determine $\omega = 2\pi/T$ from the experimentally measured T.

(ii) Using the measured maxima and minima plot of $\ln x_n$ vs t which has a slope of $\frac{\gamma}{2}$ to determine γ. Having obtained ω and γ, calculate ω_0 from equation (9).

- Critical Damping

Using Eqn. 8 and noting $x = 0$ and $\dot{x} = v_o$ at $t = 0$, one finds $x = v_o t e^{-\frac{\gamma}{2}t}$. Using $\dot{x} = 0$ at maximum deflection, determine γ from this equation using the time to maximum deflection and compare with the prediction $\gamma = 2\omega_o$.

Electrical system

- Underdamped

(i) Determine $\omega = 2\pi/T$ from the experimentally measured T.

(ii) Using the measured maxima and minima plot of $\ln x_n$ vs t which has a slope of $\frac{\gamma}{2}$ to determine γ. Having obtained ω and γ, calculate ω_0 from equation (9).

(iii) Calculate the theoretical value expected for ω_o from equation (15) using the values for L and C given on the apparatus. Determine the expected value of γ from equation (14) using the total resistance $R = R_L + R_V + R_g$.

- Critical Damping

Compare the measured value of the total resistance required for critical damping with the theoretical value. This latter is obtained from the condition $\gamma = 2\omega_o$ which is equivalent to $R^2 = 4L/C$. In practice it is difficult to distinguish critical damping from underdamped cases close to critical, e.g., the amplitude of the peak at $t_1 = T/2$ in Fig. 2 for $\gamma = 0.8 \times 2\omega_o$ is only 1% of the initial displacement at $t = 0$.

Produce a graph using the three traces saved in CSV format. Compare the trend to that seen in Figure 2.

Sample Formal Report
Analog Low Pass Filter

Abstract

The validity of an ideal circuit model for an R-C low pass filter was investigated by experimentally measuring the amplitude and phase of the filter's transfer function and comparing these measurements to the values calculated using an ideal circuit model of the filter. The experimental values were obtained by applying a 100 Hz to 100 kHz sinusoidal signal to the input of the filter and plotting the amplitude and phase of the transfer function. The theoretical and experimental values of the amplitude of the transfer function were found to agree within one standard error. The dominant source of uncertainty in the experimental amplitude transfer function (+/- 0.5 dB) was the measurement error associated with the analog oscilloscope. This was much larger than the uncertainty in the theoretical transfer function (+/- 0.05 dB) introduced by the uncertainties in the value of the resistor (R=1.473+/-0.007 KOhm) and capacitor (C= 34.2+/-0.2 nF). The theoretical and experimental measurements of the phase of the transfer function also agreed within one standard error. These results demonstrate that a simple ideal circuit model can be used to predict the response of an R-C low pass filter over the frequency range from 100 Hz to 100 kHz.

Introduction

Low pass filters are used in a variety of devices where undesirable high frequencies need to be removed. For example, telephone lines fitted with DSL splitters use low pass filters to remove the high frequency data signal from the lower frequency voice signal before it goes to the telephone [1]. Similarly, digital audio recording systems use a low pass filter to prevent signals above the audio frequency range from corrupting the audio frequency signals. A very simple analog low pass filter, referred to as an R-C filter, can be created using a single resistor and a capacitor. The input voltage signal is applied across the series combination of the resistor and capacitor and the filtered output voltage is measure across the capacitor. At low frequencies (f), the impedance of the capacitor ($Zc = 1/(j2\pi fC)$) is very large compared to the impedance of the resistor and the output voltage is approximately the same as the input voltage, while at high frequencies the impedance of the capacitor approaches zero and so the output voltage also approaches zero. The transition between when the filter passes the input signal with little attenuation and when the input signal is attenuated is determined by the cutoff frequency of the filter (fc). The cut off frequency is defined as the frequency at which the relative amplitude of the output signal with respect to the input signal is reduced by 3 dB. In addition to attenuating high frequency signals, the R-C low pass filter also introduces a phase shift between the input and output signals. At low frequencies the input and output signal are in phase but at high frequencies the output signal is shifted by -90 degrees with respect to the input signal. It is important when designing a low pass filter that the cut off frequency be placed at a low enough frequency to attenuate unwanted signals but not so low that signals in the desired frequency range experience amplitude and phase distortion. Determining the optimal cutoff frequency is usually done using an ideal circuit model [2]. The objective of this experiment was to determine how well an ideal circuit model of a low pass filter predicts the experimental frequency response of the filter.

Theory

The R-C low pass filter shown in Figure 1 can be treated as a voltage divider with impedances Z_R and Z_C.

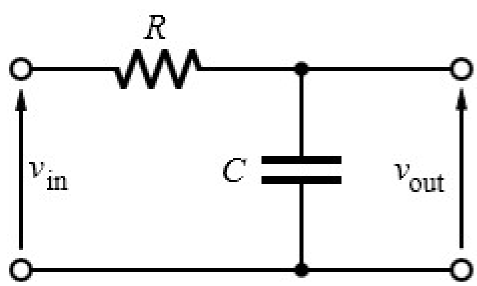


Figure 1: A capacitive first order low pass filter with resistor of resistance R and capacitor of capacitance C

From Ohm's Law we can describe the relationship between the input voltage $\mathbf{V_{in}}$ and the current through the circuit **I** in terms of the impedances:

$$\mathbf{V}_{in} = \mathbf{I}(\mathbf{Z_R} + \mathbf{Z_C}) \qquad (1)$$

or $$I = \frac{\mathbf{V_{in}}}{\mathbf{Z_R} + \mathbf{Z_C}} \qquad (2)$$

The output of the filter $\mathbf{V_{out}}$ is the voltage across the capacitor and is given by Ohm's Law as $\mathbf{V}_{out} = IZ_C$, or, using **I** from equation (2)

$$\mathbf{V_{out}} = \mathbf{V_{in}} \frac{\mathbf{Z_C}}{\mathbf{Z_C} + \mathbf{Z_R}} \qquad (3)$$

Substituting $\boldsymbol{Z_C} = \frac{1}{j\omega C}$, and $\boldsymbol{Z_R} = R$, where $\omega = 2\pi f$ we find the circuit transfer function **H** to be

$$\mathbf{H} = \frac{\mathbf{V_{out}}}{\mathbf{V_{in}}} = \frac{\frac{1}{j\omega C}}{\frac{1}{j\omega C} + R} = \frac{1}{1 + j\omega R\, C} \qquad (4)$$

The expression for the magnitude and phase of the transfer function are given in equations (5) and (6)

$$H = \left| \frac{\mathbf{V_{out}}}{\mathbf{V_{in}}} \right| = \frac{1}{\left(1 + (\omega R\, C)^2\right)^{1/2}} \qquad (5)$$

$$\phi = \angle\left(\frac{\mathbf{V_{out}}}{\mathbf{V_{in}}}\right) = \tan^{-1}\left(\frac{0}{1}\right) - \tan^{-1}\left(\frac{\omega RC}{1}\right) = -\tan^{-1}(\omega R\, C) \qquad (6)$$

The experimental value for the amplitude of the transfer function is simply

$$H = v_{out}/v_{in} \qquad (7)$$

And the phase can be determined from the time shift Δt between the input and output voltages as

$$\phi = 2\pi \frac{\Delta t}{T} \qquad (8)$$

The cut off frequency of the filter is defined by the frequency when the amplitude of the output signal is reduced by 3 dB with respect to the input signal, $20\log(v_{out}/v_{in}) = -3\text{dB}$. [3] This will occur when $v_{out} = \frac{1}{\sqrt{2}} v_{in}$.

Using this voltage ratio in equation (1) we find $\omega RC = 1$ so the cutoff frequency for the circuit in Figure 1 is

$$f_c = \frac{1}{2\pi RC} \qquad (9)$$

From equation (6) we find that that the phase shift at the cutoff frequency is expected to be -45^0. These equations assume that all elements are ideal, the capacitor has no resistance, the resistor has no capacitance or inductance, and the leads have no inductance or resistance.

Apparatus and Experimental Procedure

The low pass filter shown in Figure 1 was constructed on a breadboard using a 1.473 kOhm resistor and a 34.2nF capacitor of. The resistance and capacitance were measured using a Wavetech CR50 multimeter which has an uncertainty of 0.5%$\pm$1digit. Figure 2 shows the circuit used to measure than filter transfer function. A BK Precision 4012A signal generator was used to generate an input sinusoidal voltages ranging in frequency from 100 Hz to 100 kHz. Seven different frequencies were selected over this frequency range so that the measurements would be approximately equally spaced when plotted using a log frequency scale. At each frequency, the amplitude and time shift between the input and output voltage waveforms were recorded using an analogue oscilloscope (Hitachi V252). The input impedance of the oscilloscope was 1 MΩ, which is much larger than the impedance of the resistor and capacitor and could therefore be ignored in the analysis. The ground for the oscilloscope leads were both connected to ground on the breadboard. The amplitude and time measurements were made by using a visual inspection of the oscilloscope screen. The uncertainty was taken as half the smallest tick mark (there were 5 ticks per division) for all measurements. The uncertainty varied as the amplitude and time base scales were changed to obtain the best display the waveforms.

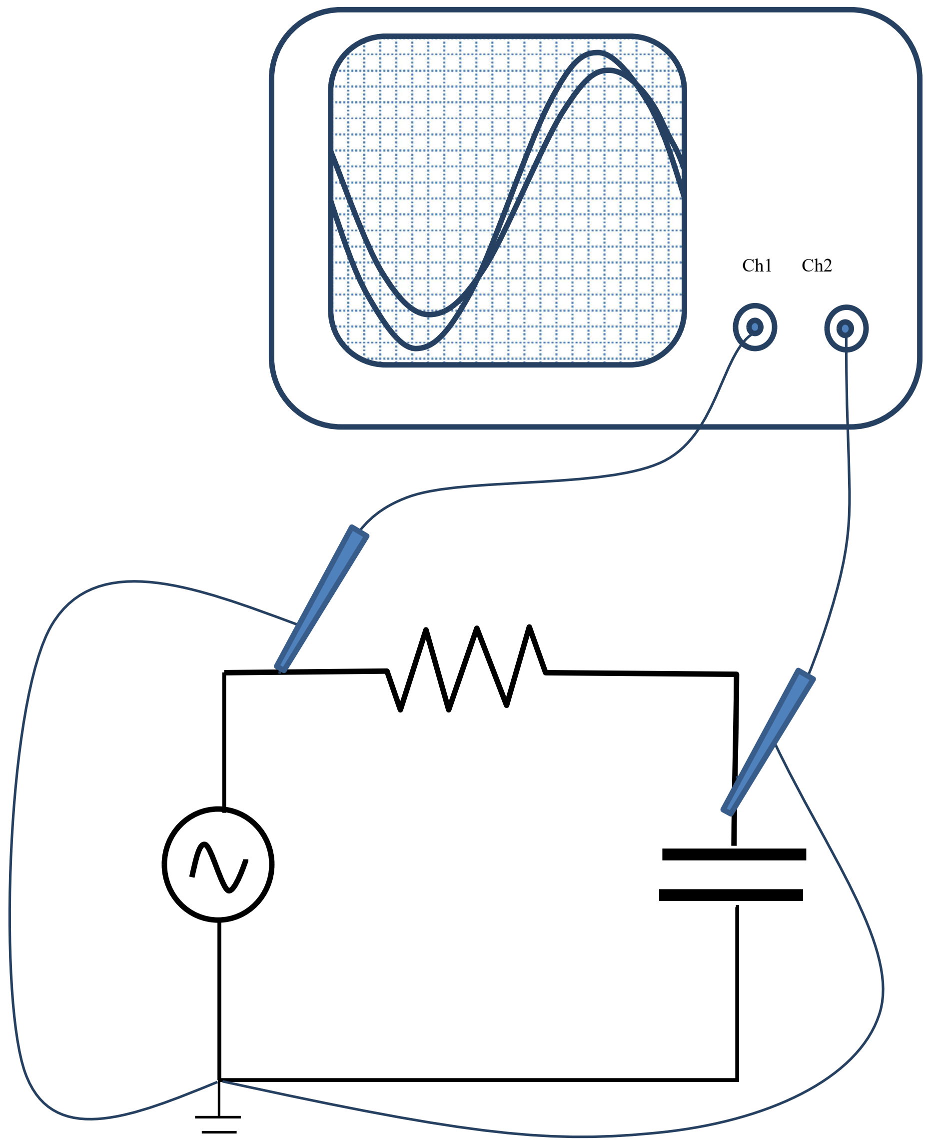


Figure 2: Low pass filter with an oscilloscope set up to measure the input (ch1) and the output (ch2).

Results and Analysis

The input and output voltages for the low pass filter were used in equation (7) to calculate the magnitude of the experimental transfer function *H* for the frequencies tested. These points were plotted in Figure 3 along with the expected magnitude from equation (5). The experimental points all fall on the expected line within uncertainty. The cutoff frequency calculated from equation (9) is 3160 $\pm$40 Hz. This is also seen in Figure 3 and matches well with the -3dB point of the experimental data.

The time shift data was used to determine the experimental phase shift from equation (8) and was plotted along with the expected curve from equation (6) and the cutoff frequency in Figure 4. All experimental data points fall on the line of the expected phase shift. The phase shift at the cutoff frequency is approximately 45^0 as expected from equation (6).

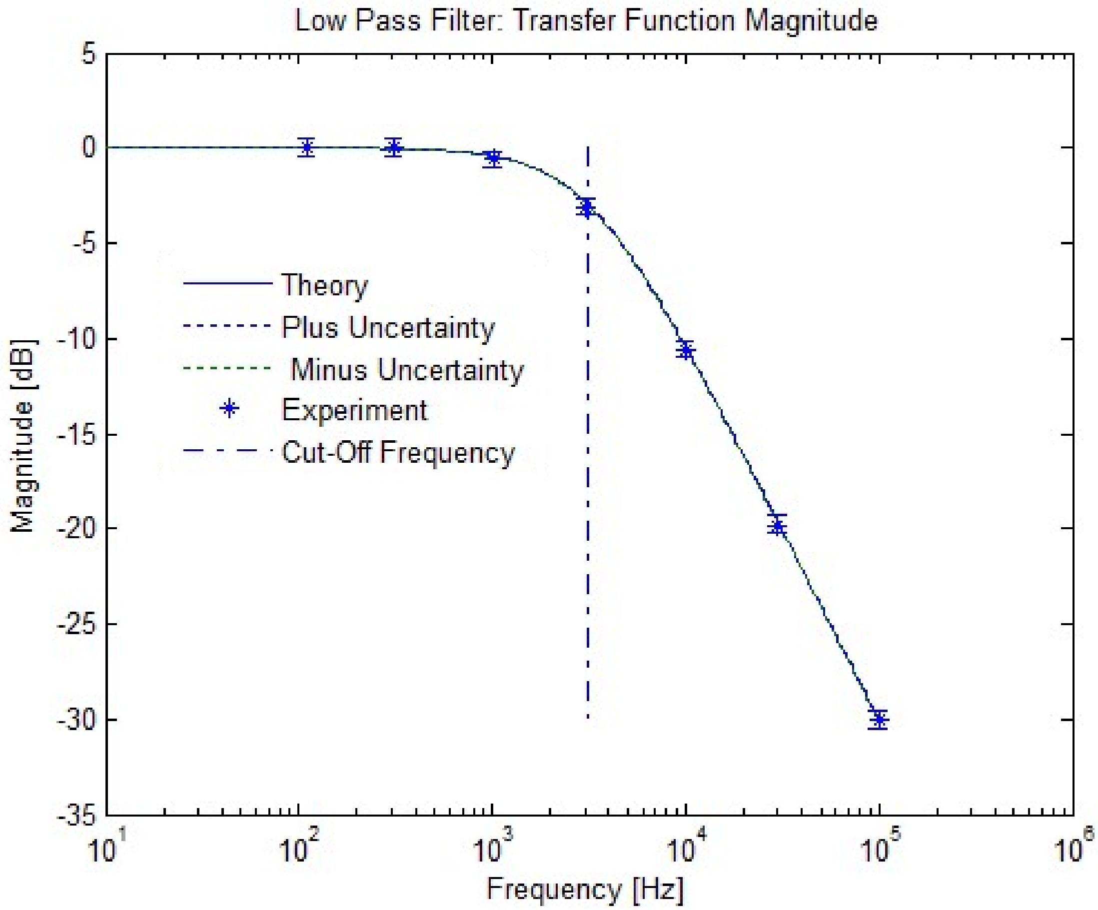


Figure 3: Low pass filter transfer function magnitude – experimental points and theoretical curve. R=1.473 kHz C= 34.2nF . Error bars display the uncertainty in the experimental points. Uncertainties on theoretical curve are shown as dashed lines but are not visible because the uncertainties are very small.

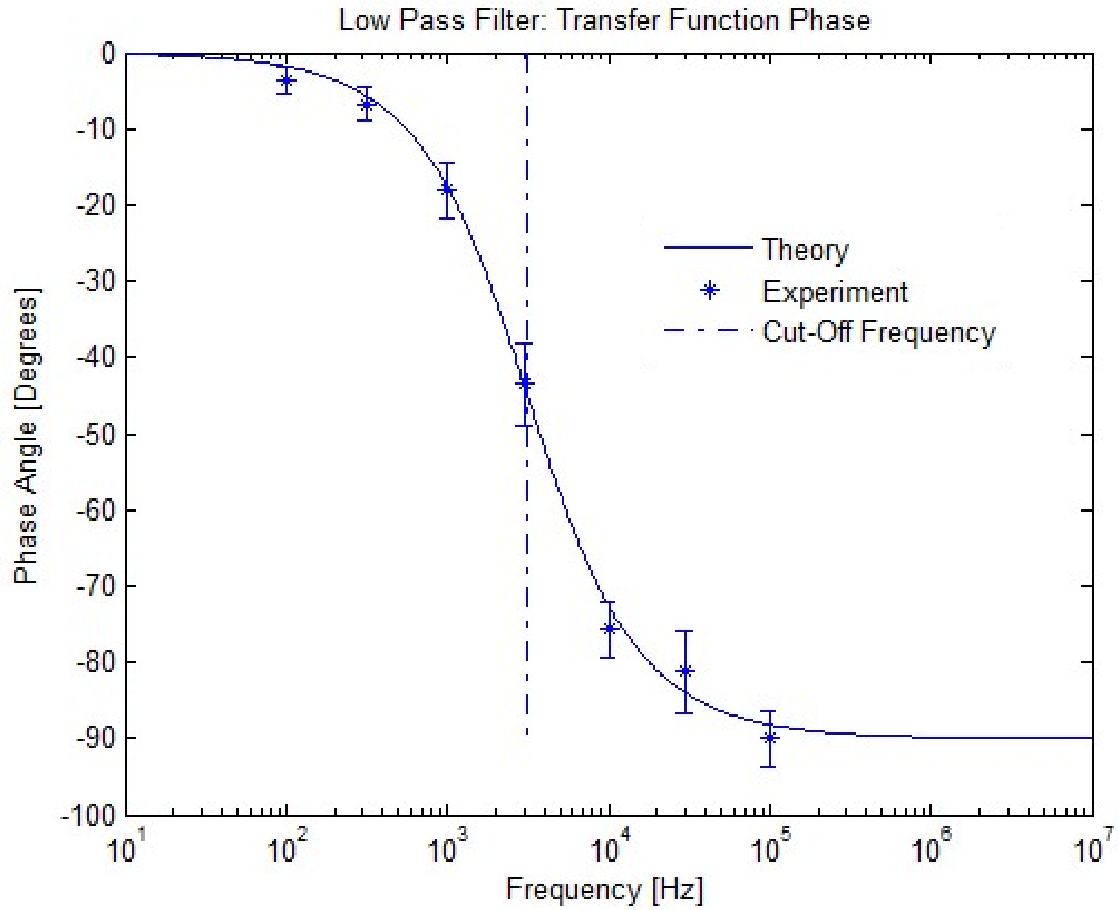


Figure 3: Low pass filter transfer function phase – experimental points and theoretical curve. R=1.473 kΩ C= 34.2nF .

Discussion

The magnitudes of the experimental and theoretical transfer functions for the low pass filter were found to be very similar. Both curves start at a magnitude of 0 dB and show little change until the cutoff frequency. At the cutoff frequency the magnitude of the transfer function was close to -3dB. At frequencies above the cutoff frequency the magnitude of the transfer function decreased at a constant rate with a slope of -20dB/decade.

The experimental and theoretical curves for the phase of the transfer function were also very similar. For both curves the input and output signals were in phase until close to the cutoff frequency, about -45^0 at the cutoff frequency, and then the phase of the transfer function rapidly decreased to -90^0.

A low pass filter is used in a circuit to attenuate signals at frequencies above the cutoff frequency while passing signals below the cutoff frequency. The plots of the magnitude and phase of the transfer function show that the low pass filter was functioning as expected. The amplitude of the output signal was identical to the input signal for frequencies below the cutoff frequency but attenuated at frequencies above the cutoff frequency. Similarly, there was no phase shift between the input and output signals for frequencies below the cutoff frequency. This indicates that signals below the cutoff frequency would be passed with no distortion.

The theoretical cutoff frequency was calculated as 3160 $\pm$40 Hz and was in good agreement with the experimental data as seen on the graphs. The experimental data for both the magnitude and phase of the transfer function agreed within the uncertainty of the theoretical curves.

The excellent agreement between the theory and experiment indicate that it was valid to treat the resistor and capacitor as ideal circuit elements in the theoretical model of the filter. There did not appear to be any oversimplifications in the theoretical model. The excellent agreement also indicates that the design of the experiment was appropriate and that equipment used was suitable to measure the experimental transfer function.

The uncertainty in the theoretical transfer function was due to uncertainties in the measured values of the resistor and capacitor. Since these values could be measured very accurately using a digital multimeter, the resulting uncertainty in theoretical transfer function was very small compared to the uncertainty in the experimental transfer function. For example, the uncertainty in the theoretical magnitude of the transfer function was approximately +/- 0.05 dB while the uncertainty in the experimental magnitude values was approximately +/- 0.5 dB. The experimental amplitude and time measurements for this experiment were taken through a visual analysis of the display on the oscilloscope. This represents the greatest source of random error found in the experiment. Uncertainties in the amplitudes and time shifts were assumed to be half of the smallest division. A digital scope that can more accurately measure distance between zero crossings (the time shift) would improve the accuracy of the phase measurements. As well such a scope could measure the amplitude of the signals more accurately.

Conclusions

A simple, first order capacitive low pass filter was built on a breadboard and measurements of the amplitude of the input and output as well as the time shift between them were made using an oscilloscope so that the magnitude and phase of the filter`s transfer function could be determined experimentally. The experimental transfer function was compared to the theoretical transfer function and was found to be in good agreement. The assumption of ideal elements was appropriate and resulted in a valid model. The test set up, while relying on visual inspection of the analogue display, still produced accurate results. A digital scope which allows magnitude and phase to be measured more quickly and accurately would make the experiment easier to run and further improve the accuracy of the experimental transfer function.

References

[1] http://en.wikipedia.org/wiki/Low-pass_filter, Accessed May 30, 2012.

[2] http://www.allaboutcircuits.com/vol_2/chpt_8/2.html, Accessed May 30, 2012.

[3] P. Horowitz and W. Hill, The Art of Electronics, Cambridge: Cambridge University Press, 1989.

Appendix – Raw Data and Sample Calculations

• • •