Physics 25L lab manual

David Stuart, UCSB Physics Department

January 6, 2015

Contents

1 Introduction to experimental methods

Experiment is the foundation of physics, indeed all of science. The primary goal of this course is to teach you experimental methods and indeed the experimental mindset. This section provides an overview of three critical aspects of experimental work: rigorous record keeping, thorough validation, and proper determination of uncertainties. To succeed in the lab, and *in your future research endeavors*, these need to become instinctive. You should have already had some of this drilled into you from what you've learned so far in solving physics problems, and I'll use analogies to that to introduce these concepts and give specifics of each.

1.1 Uncertainties

One of the first topics discussed in introductory Physics is the concept of significant digits. If you are asked to calculate the acceleration of an object with mass $m = 33$ kg acted on by a force of $F = 101$ N, you know to say $a = 3.1\ \mathrm{m/s^2}$ rather than $a = 3.060\overline{606}\ \mathrm{m/s^2}$, right? If the mass is only given with two significant digits, then you only know the acceleration to a precision of two significant digits. That is the concept that underlies experimental uncertainties; we need to take into account the precision of what we observe.

Rather than encoding the precision in the number of significant digits, which is ambiguous, we make it explicit by specifying an uncertainty. For example, we might state that the mass is $m = 33 \pm 0.5$ kg, where the 33 is called the central value and the 0.5 is called the uncertainty. It is also called the "error", but that should not be confused to mean that it is a "mistake"; the word "uncertainty" is a more direct description. You'll end up using both, simply because the term "error" is quick and easy.

When we say that $m = 33 \pm 0.5$ kg, we don't mean that we know with full certainty that the mass is somewhere between 32.5 and 33.5 kg. Not much is known with full certainty. Rather, the 0.5 quantifies the uncertainty. Indeed, if properly constructed, it should correspond to us being 68% confident that the true value of the mass lies between 32.5 and 33.5 kg. The choice of 68% confidence comes from the normal (or Gaussian) probability distribution, which describes most real-valued random variables. The simple idea is that when we measure a mass, what we obtain is the true value perturbed by a set of random influences. Those random influences are just a bunch of physics processes that affect the measurement in difficult to quantify ways, e.g., a breeze, or vibrations of the measuring device, or simply how well we can see where the pointer is on a scale. Since the direction of their contribution is random, they'll have the effect of randomly changing the observed value according to a probability distribution.

The Gaussian distribution is a good estimate for that probability distribution for several reasons. A particularly important one is that, because of the central limit theorem, the combination of a large number of random variations approaches the Gaussian distribution. So, combining a large number of measurements gives a distribution that approaches a Gaussian. The Gaussian distribution is

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

where μ is the mean of the distribution and σ is called the standard deviation. The probability of finding x between $\mu - 1\sigma$ and $\mu + 1\sigma$ is the previously mentioned 68%, which is why we target the uncertainty we quote for our measurements to correspond to a 68% probability range.

Once you have that 68% definition for a 1σ difference from μ, you can determine the probability for any other difference. The probability for $|x - \mu| > 2\sigma$ is 95%, and the probability for $|x - \mu| > 3\sigma$ is 99.7%. Knowing how the variation of measurements should spread is very useful for making sure that the uncertainties are properly determined, as discussed below.

1.1.1 Estimating uncertainties

Every time you make a measurement, you also need to determine the uncertainty. Sometimes that can be calculated, but it often requires an estimate. A common situation is where you read a number on a display, such as a digital scale used to measure a mass. The scale will read some number of digits, such as $m = 33.0$ kg. When the reading is limited by the granularity of the display, you should specify the uncertainty to be the size of the least significant digit divided by $\sqrt{12}$. The idea behind this rule of thumb is that the round-off error in the last digit is 0.1, i.e., somewhere between 32.5 and 33.5. So, if you do this for measurements of many different masses, the true value will be uniformly distributed between the round-off range. And the 68% probability range for that uniform distribution is $1/\sqrt{12}$. Each individual measurement won't be Gaussian distributed; rather it is uniformly distributed. But a large number of measurements will be Gaussian distributed due to the central limit theorem.

This $1/\sqrt{12}$ trick also works when reading where a pointer is on a dial or when determining where a point is on a ruler; you can determine a range of values such that you are nearly 100% confident the true value lies between them. Then the span of that range over $\sqrt{12}$ is a good estimate of the 1σ uncertainty.

With this approach, you can quickly estimate an uncertainty for each and every value that you measure. And you should. In your lab work, *every number* should have an uncertainty quoted with it. Write $m = 33.00 \pm 0.03$ kg, not just $m = 33$ kg.

But, what about significant digits? The convention for that is to write the uncertainty with 1 significant digit and then write the central value with a matching number of digits. That is what is done in the 33.00 ± 0.03 above. The reason to use only one significant digit is that precision on the uncertainty is not material to the result. The conclusions about the probability don't really change if the standard deviation is 0.034 or 0.0025 instead of 0.03, so you can round away the second digit. A nuance to this is that there is a larger difference between a standard deviation of 0.1 and 0.14 or 0.05, so it is worth keeping two significant digits when the least significant digit in the uncertainty is a 1. (Many physicists dispense with that nuance and just always use two significant digits for the uncertainty. That is reasonable, but the significant digits on the central value should always match.)

1.1.2 Propagating uncertainties

You will often measure some parameters and then use them to calculate another parameter, so you'll need to propagate the uncertainty from the original parameters to the calculated one. An example is measuring the length of an object; to do that you measure the position of its two ends, x_1 and x_2 and calculate the length as $L = x_2 - x_1$. The uncertainty on L is larger than the uncertainty on either x_2 or x_1 because deviations present in each combine. In any single measurement, though, the deviation might cancel a bit or they might add; the direction of one with respect to the other is random. This is accounted for by adding the two uncertainties "in quadrature", so $\delta L = \sqrt{\delta x_1^2 + \delta x_2^2}$, where I've introduced the notation of δX to represent the uncertainty on X. (The same quadrature sum occurs if you were to add x_1 and x_2 rather than subtract them.)

If your calculation involves multiplying or dividing two numbers, such as either $C = A * B$ or $C = A/B$, then the fractional uncertainties add in quadrature, i.e., in both cases the uncertainty on C can be found from $\delta C/C = \sqrt{(\delta A/A)^2 + (\delta B/B)^2}$.

These two are just special cases of the more general formula, which is easy to state and explain. If y is a function of several variables, x_1, x_2, ..., x_N, then the uncertainties on the x_i values propagate to an uncertainty on y obtained from

$$\delta y = \sqrt{\sum_{i=1}^{N} (\frac{\partial y}{\partial x_i} \delta x_i)^2} \tag{1}$$

This is easy to remember if you just think of $\partial y / \partial x_i$ as the slope of the multidimensional y function in the x_i direction. We multiply that slope by the step size in the x_i direction, δx_i, to get x_i's contribution to the total uncertainty.

As an example, suppose that you wanted to measure the acceleration due to gravity using a simple pendulum. You could measure the length of your pendulum, L, and its period of oscillation T and then calculate g from $g = 4\pi^2 L/T^2$. To propagate the uncertainties L and T to an uncertainty on g we'd use

$$\delta g = \sqrt{(\frac{\partial g}{\partial L}\delta L)^2 + (\frac{\partial g}{\partial T}\delta T)^2} = \sqrt{(\frac{4\pi^2}{T^2}\delta L)^2 + (\frac{-8\pi^2 L}{T^3}\delta T)^2}$$

That is a bit ugly, but notice that you can rewrite it as

$$\delta g = \sqrt{(\frac{4\pi^2 L}{T^2}\frac{\delta L}{L})^2 + 4(\frac{4\pi^2 L}{T^2}\frac{\delta T}{T})^2} = \sqrt{(g\frac{\delta L}{L})^2 + 4(g\frac{\delta T}{T})^2} = g\sqrt{(\frac{\delta L}{L})^2 + 4(\frac{\delta T}{T})^2}$$

So, we can write the fractional uncertainty, $\delta g/g$, as

$$\frac{\delta g}{g} = \sqrt{(\frac{\delta L}{L})^2 + 4(\frac{\delta T}{T})^2}$$

It is a quadrature sum of the fractional uncertainties on L and T, but the contribution from T is weighted by a factor of 4 because T comes into the calculation as a square.

While planning how you'll measure something, you should use a calculation like this to determine how important each observable is in the total uncertainty. In this case, we should spend more time planning how to get a relatively precise measurement of T than of L because of the factor of 4.

1.1.3 Plotting and fitting

As noted above, when you write down a measured value, it should always have its associated uncertainty included. Similarly, when you plot measured values, they should always include a representation of the uncertainties. This is usually done by putting bars on data points, which are called "error bars". An example is shown in Figure 1 below. It represents 100 repeated measurements of the period of the pendulum, plotted vs the measurement number. The vertical error bars on the points show the uncertainty on each measurement, which we estimate to be $\delta T = 0.03$ s. Sometimes, plots like this could have a horizonal error bar if there is also an uncertainty on the value plotted on the x-axis, but there is no uncertainty on the measurement number plotted here.

The plot includes a horizontal line that corresponds to the average of all the points. (It is actually a fit, as discussed below). The uncertainty represented by the error bars should represent a 68% probability, so about 68% of the points should have their error bars touching that average line, and 32% of them should be more than 1σ away and thus not have their error bars touch the line. About 5% of the points should be at least 2σ away. By looking at the scatter, we can determine if the measured values, and importantly their uncertainties, are properly determined. The points in Figure 1(a) look reasonable, but we want to be more quantitative.

You can quantitatively determine if the scatter of the measurements is proper in two ways. First, we can simply plot the distribution of measured values. This is done in Figure 1(b), which makes

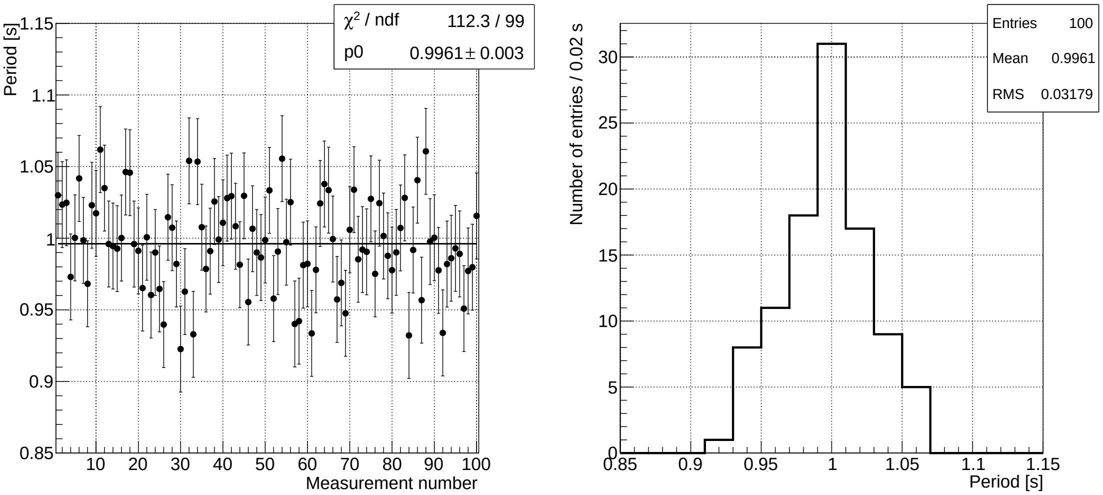


Figure 1: An example of plotting measurements, which is the result of 100 repeated measurements of the period of a pendulum. In (a) we show the measured period vs the measurement number (i.e., the time of the measurement). In (b) we *histogram* the measured values to see their distribution. It is distributed like a Gaussian with a standard deviation (obtained from the RMS) of 0.03 s. See the text for a discussion of fitting and χ^2.

a "histogram" of the measured values, i.e., it splits the range of measured values into fifteen "bins" between 0.85 and 1.15 seconds. Then on the y-axis it plots the number of measurements that have a value falling within each of the bins. The result should be a Gaussian distribution where the mean, which is the average of all the measurements, is the best estimate of the true value, and the standard deviation corresponds to the uncertainty on any single measurement. The standard deviation is shown by the "RMS" number in the upper right corner. (RMS is an acronym for "root mean squared", which is the procedure used to calculate the standard deviation, so $\mathrm{RMS} \approx \sigma$). The fact that it equals 0.03 confirms the estimated uncertainty of 0.03.

The uncertainty on the mean is much less than 0.03 because we can combine all the measurements together to obtain a more precise result. If all the measurements have the same uncertainty, as in this case, the average has an uncertainty that is reduced by a factor of $1/\sqrt{N}$ from the σ of a single measurement. (In it is a useful trick to remember that the uncertainty on the mean of a Gaussian such as the one in Figure 1(b) is $\approx \mathrm{RMS}/\sqrt{N}$, where N is the number of entries.)

If the uncertainties on each of the measurements are not all the same, then we need to use a "weighted average", which is

$$\overline{T} = \sum_{i=1}^{N} \frac{T_i}{\sigma_i^2} / \sum_{i=1}^{N} \frac{1}{\sigma_i^2}. \tag{2}$$

We can also "fit" the data to find the best central value. This is often done with a χ^2 minimization, where we adjust the parameters of some function $f(x)$ to obtain the parameter values that best fit the measurements. In general, the chi-squared is

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{y_i - f(x_i)}{\sigma_i} \right]^2, \tag{3}$$

where y_i is the value of the i-th measurement, $f(x_i)$ is the prediction for y_i based on the function operating on the x-value for the i-th measurement, and σ_i is the uncertainty on the i-th measurement. With this definition, χ^2 corresponds to the summed deviation between the measured values, the $y's$, and the predicted values, the $f(x)'s$, where the deviations are weighted by their individual precisions. The parameters for the function f that minimize that total deviation are the “fitted values”.

In the example of averaging a number of measurements, the function that we want to fit is just a constant, so $y_i = T_i$, and $f(x_i) = \overline{T}$ is just the average we are trying to extract. (We’ll see a less trivial example of a fit below). In Figure 1(a), the horizontal line is the result of such a fit to a constant function (i.e., a 0-th order polynomial), and the box in the upper right hand corner shows that the result of the fit is $p0 = 0.996 \pm 0.003$. The name “p0” just means the 0-th order coefficient of the fitted polynominal; in our case this is $\overline{T}$.

Note also that the resulting χ^2 value is given in the plot; $\chi^2/\mathrm{ndf} = 112.3/99$. That means that the value of the χ^2 at the fitted minimum is 112.3, and ndf represents the “number of degrees of freedom”, i.e., how many data points are included in the fit (100 in our case) minus the number of function parameters that are allowed to vary (1 in our case). For a reasonable fit, the χ^2 should be about equal to the number of degrees of freedom. If it is much less, then the uncertainties are probably overestimated. If it is much more, than either the uncertainties are underestimated or the chosen functional form is incorrect. The value in this case, is reasonable. (At http://plot.physics.ucsb.edu/StatCalc/ there are several online statistical tools including one to calculate the probability of obtaining $\geq$ a certain χ^2 value for a given number of degrees of freedom. Our probability value, also called the “p-value”, corresponds to 18%, which is quite reasonable.)

But, suppose our repeated measurement of the pendulum’s period resulted in the data shown in Figure 2(a). Now, if we fit the data to a constant it is an obviously bad fit; the first half of the data points are below the average and the second half are above it. You can also see that the χ^2 value is anomalous. The probability of obtaining a $\chi^2 \geq 483.7$ with 99 degrees of freedom is miniscule, 5×10^{-52}, so the data are not consistent with being independent of time (for which the measurement number is a proxy). But, if we change the fit function from a constant to a line, then the fit looks good, as shown in Figure 2(b). And, the fit returns a slope for the line given by the 1st order coefficient, $p1 = 0.0020 \pm 0.0001$. Note the number of significant figures used is as previously discussed, but the leading zeros are distracting. It is better to use units that make the numbers easier to read, such as 2.0 ± 0.1 ms/step.

This example of repeatedly measuring the period of a pendulum and fitting the results as a function of

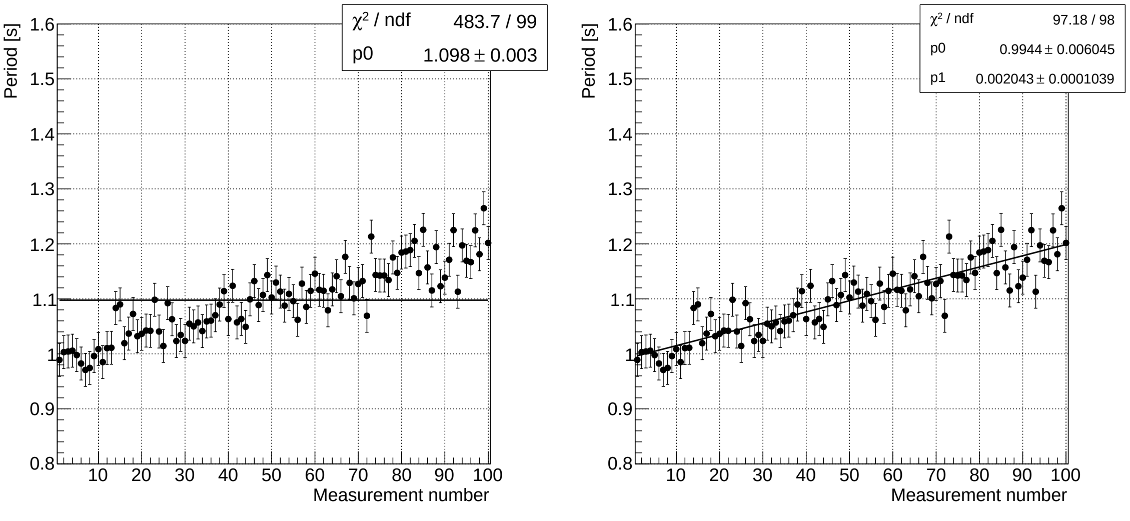


Figure 2: A second example, where we plot the measured period of a pendulum vs the number of (repeated) measurements. In (a) there is a significant dependence on the time of measurement; the data is not well fit by a constant (as shown by the line). In (b) we instead fit that data to a line and find that it has a significant slope. See the text for a discussion of fitting and χ^2.

the measurement number is pretty simple. It can be useful in validating that the measurements are stable over time (as discussed in the next section). A more likely scenario is that you measure one parameter as a function of another, such as the displacement of a spring as a function of the mass you hang on it. Then, you can fit the data points to a function to extract a parameter of interest. In this case it would be $x(m) = \frac{g}{k}m$, and you could determine the spring's force constant from the slope of the fitted line.

Such use of fitting to analyse multiple measurements with varying parameters is much more powerful than simply calculating numbers from a single measurement. You should use it whenever you can during this course. There are many different data analysis software packages available, for example Mathematica, MATLAB, NumPy, and ROOT. These packages take a while to learn, but it pays off down the road to pick one that you want to become familiar with and use it. (Accounting software such as Excel isn't useful for this because it doesn't handle uncertainties well.) To get you started making plots without a steep learning curve, however, an online plotter is available at http://plot.physics.ucsb.edu/ which gives a web interface to the ROOT analysis software and will allow you to enter, plot, and fit data. Feel free to use it in this course, or your own preferred analysis, plotting, and fitting tool.

1.2 Validation of results

When you finish a physics problem, you check your answer, right? Sometimes answers are given in the back of the book, but for real research problems, the answers are not given or even known. The first time a problem is solved, or a quantity is measured, the physicist doing that research had to check it somehow. There are several ways to check, or validate, a result.

Checking to see if your measurement matches the "accepted value" is **not** one of them! If students measure g with a pendulum, for example, it is a common temptation to think that if they get $9.8\ m/s^2$ they did it right, and if not, then they did it wrong. Some even incorrectly call the deviation between their measured value and the "accepted value" the "error". Don't be that guy. If you are measuring g with a pendulum imagine that you were the first person to ever do so. Your measurement will then *be* the value. You might say that you aren't the first person to do so, but you are preparing to do research, where that will be the case. (That is the whole point of research). To continue with the example of measuring g, there are two scenarios where this oft repeated measurement can break new ground. If you were the first person to step on a new planet, you could pull a yo-yo out of your pocket and, in one minute, know the mass of the planet.[1] As a more down to earth application, precision maps of the variation of g across the surface of the earth provide a useful way to find underground oil reserves. There, the payoff comes from measurements that *don't agree* with the standard value of g.

So you want to make sure that your result is valid without checking it against someone else's answer. You've already learned several ways to validate your answer when solving physics problems: you can check that the units are correct; you can see if the limiting behavior is physically meaningful; and you can plug in a few numbers to see if the results are reasonable. For experimental work, there are also tricks to use. I've listed some of them below.

- **Measure your uncertainties**
  Whenever possible, find a way to take simple repeated measurements of the basic ingredients and check if the spread of resulting values is consistent with your estimate of the uncertainties, like we did in Figure 1(a). If the spread doesn't match, you have learned that your uncertainty estimate is wrong for some reason.

---

[1] If you argue that you would not have safely landed without already knowing that, then I can only say touché.

- **Check your raw data**
  Sometimes you will make a mistake, i.e., the data really included an *error* as opposed to just a random uncertainty (which we sometimes colloquially call "errors", as noted above). You can often find and correct these mistakes with sanity checks on your raw data.

  A good example of this is checking for time dependence. Some things are supposed to be time dependent, but others are not. If in Figure 1(a) you found that the 37th measurement deviated by 7σ from the average of the others, you'd be justified to reason that you made a mistake in that measurement and remove it from the final measurement. Don't erase it from your logbook; in fact you should carefully document why you decided to remove it. (This will happen ocassionally, especially if you are writing down or typoing in numbers).

  You might find, as in the example of Figure 2, that the period of the pendulum increases slowly over the repeated measurements. That means that there is a systematic drift over time in some aspect of your measurement. Now, g might be drifiting over time. You have to decide if you expect that to be reasonable, but Occam's razor motivates you to look for a simpler explanation. The string that you use for your pendulum could be stretching. So, this sort of raw data validation would motivate you to redesign your technique so as to measure the string each time and correct for any stretching.

  What checks should you do? The answer is "Every one that you can think of and carry out in a reasonable manner". Often, this can be done by recording as many parameters about the measurement as possible. You may not expect the string to stretch, but you should remeasure it each time anyway. You don't expect g to depend on temperature, but you should record the temperature anyway. Once you have all of your data, you can then plot measured values such as the period, and extracted values such as g, as a function of the other parameters you've measured. Any significant dependencies tell you something about the validity of your measurement.

- **Repeat the measurement in systematically different ways**
  It is a basic tenet of science that a discovery will only be considered valid after it is confirmed by an independent measurement. For example, as exciting as the Cold Fusion claims in 1989 where, they were not replicated and could not be validated. But, you shouldn't just wait and hope that someone else repeats your measurement and confirms the result. You should repeat it yourself. That doesn't mean someone else won't have to confirm your discovery for it to be recognized, but it will help you avoid the pitfalls that lead to results that are later shown to be invalid when others fail to replicate them.

  So, repeat your measurements with varied setups. The variations allow you to tell if there are unexpected dependencies or underestimated uncertainties. Suppose for example that you measure g by recording the period, T, of a pendulum of length L and using $g = 4\pi^2 L/T^2$. For a given stop watch precision, i.e., a given δT, you'll obtain the most precise value of g by using as large an L as possible. But, to test for systematic effects, i.e., to validate the measurement, it is best to measure g with several different lengths. The shorter lengths give a less precise result, but combining the different lengths still provides improved precision over a single measurement. And, more importantly, the ensemble of measurements gives the most robust result. Fit the values of T^2 measured at different L to a line. The slope is proportional to $1/g$, and the goodness of fit is a validation of the measurement.

  In addition to measuring with varied setups, you should plan to make measurements in multiple independent ways. For example, you can measure the force constant of a spring by using $F =$

kx or by using the fact that the period of its oscillation is proportional to $\sqrt{k/m}$. These two approaches have different systematic uncertainties, and how well they agree is validation of the measurement.

You will need to get into the habit of thoroughly validating your measurements. Your success in research will depend on that. More immediately, but less importantly, your grade in this class will also depend on it. Don't just measure something once! Measure dependencies. Show plots of these dependencies in your logbook, and use a fit of that dependence to extract the measured parameter you are after.

1.3 Rigorous record keeping

You should have learned to setup and write out your problem solutions in an organized and well documented way. This is particularly important for solving lengthy and difficult problems. If you don't draw a diagram and clearly state your coordinate system and assumptions, you could easily confuse yourself when you come back to the problem the next day and don't remember the details. Research is about solving difficult problems, the ones that have not yet been solved; that takes time. You might spend a year on an undergraduate research project, and Ph.D. research projects typically take several years. If you have lunch at the nano-cafe, look around at the graduate students. You'll notice that many of them have a logbook with them. This is usually a spiral bound book, although e-logs on a laptop or tablet are taking over. Why bring it to lunch? Because lunch is a good opportunity to talk over the results and plans with your colleagues, but not from memory!

You can't succeed with a long project without a carefully organized record of your work. To help you learn that approach, this course will require you to keep a rigorous record of your work in a logbook. Indeed, much of the grade will be based on the depth and quality of your logbook. Specific requirements of "logbook etiquette" are listed below:

- **Use a bound notebook that contains all your work on a given project**
  You can't possibly be organized if you just use scattered sheets of paper! In the case of an e-log, it is still important to keep the files "bound" together to keep things coherent and easy to find.

- **Don't erase**
  In fact, use a pen. If you find that something is incorrect, just put a line through it and jot down why you think it is wrong. A **thin** line is best so that you can still read it; you might actually find out later that you were wrong about it being wrong. In the case of an electronic logbook, "don't erase" becomes "don't delete". I use an electronic logbook, and it has many entries with a note that say something like "This measurement is invalid; see Entry #394 for an update". That is just an electronic equivalent of pointing to a page with the corrected version. I keep the invalid entries, and often find them useful to remember what I did – both what I did it and how I figured out that it was wrong.

- **Document obsessively**
  You should include as much information as possible in your logbook. You should work out the relevant theory *within* your logbook. Don't just do it on scratch paper and copy in the final equation; do the work in your logbook. It may not be neat, but it should be complete. You should also include diagrams and photos of your apparatus. And, you should include the raw data. If you are writing down measurements that you will analyze, write them *as a table in the logbook*. If you record things electronically, rather than in writing, then write the data's filename in your logbook.

In fact, it is best to keep all your files in a common place so that they are easy to find later based on a short descriptor within your logbook, and so they are easy to backup. I have backup disks in three locations – across two continents – with archives of my > 40k logbook entries and associated data. An earthquake may destroy one of the backups, but I can still look up the raw data for that spring constant measurement of my youth. Yes, be that obsessive about record keeping!

- **Make specific conclusions**
  You should have lots of detail in your logbook, but you'll want to quickly skim over previous work to look up the key conclusion. So, write frequent mini-conclusions in your logbook and highlight them to be easily found. For example, you might spend an hour, and a few pages, calibrating an instrument such as a stop watch. Your logbook should have the raw data and a plot showing the results, and then there should be a one line, highlighted, statement like: "I found that the stop watch time is correct to within 0.06 second." That is the key number that you'll want to quickly find and use later.

  You should also write a conclusion at the end of each time period. If you spend an afternoon in the lab, write a few lines of conclusion at the end of the day saying what you accomplished and what you think the next steps are. The next afternoon's work should begin by rereading that conclusion. You may change your mind about what the next steps are, but at least you will remember what you were thinking.

  You should write a short conclusion at the end of each measurement. For this class, that means either each lab or your multi-week individual measurement. This should be about a half page summarizing what you did and what you found. It should not be touchy feely blathering about how much you learned or how cool lasers are. It should be just the facts. For example, "I measured the frequency of oscillation of pendula with several lengths, $L = 0.5, 1.0, 1.5,$ and 2.0 m, to determine that the acceleration due to gravity in my bedroom is $g_{\text{local}} = 11.2 \pm 0.1\ \mathrm{m/s^2}$. The results obtained with the different lengths were consistent within 5%. The measured value is significantly discrepant from the expected value of g; an explanation for the deviation is not known." (Although it might be tempting to speculate about the reason for the deviation, such as the mass of your new neighbor, that speculation doesn't belong in the conclusion. Keep it to what you know.)

- **Keep it organized and timestamped**
  You should write into your logbook in chronological order; don't leave space to be filled in later. And write the date and time of each entry. There should be at least one timestamp on each page. This will, of course, mean that the logical thread may skip around when you read it later. That is OK, and it is expected. You can help keep the threads organized by putting a table of contents on the first 3 pages, which you fill in as you go. You can also jot down pointers from one section to the next logically connected section, e.g., at the bottom of page 18 you can later write, "See page 47 for a continuation of this measurement using a more precise spectrometer."

Proper record keeping is important for successful and coherent research. Your grade in this course will depend mostly on how well you implement this in your logbook record of your measurements.

1.4 Excelling

Finally, I have two suggestions for how to excel in this class. For each of the lab's, the manual has a pre-lab section that gives an introduction. Read it the night before coming to lab; it may have suggestions for what you should prepare or bring with you. The manual also has suggestions for "Going Beyond" at the end of each lab section. Read it and think about what novel thing you can do. Please wow me; it will help me recommend you for research opportunities. Writing a recommendation letter is like reporting my measurement of your skills. Stating the uncertainty on any measurement is important, and that requires having a lot of data.

2 Gamma ray absorption

Abstract: In this lab we will measure the absorption of gamma rays in matter.

2.1 Pre-lab preparation

Just as the electrons in an atom can transition between energy levels by emitting photons of specific energies, the nucleus of an atom can also transition between energy levels. However, the energy levels in the nucleus are typically in the MeV rather than eV range, so the photons emitted are quite energetic. These energetic photons are called gamma rays, or just γ's for short. They are emitted with specific energies just like the atomic transitions, and measuring the spectrum of gamma ray energies can be used to identify the composition of a sample.

A *radioactive* element is simply one with a nucleus that is in an excited state. Decays of this unstable state emit radiation, such as γ's, and leave the nucleus in a lower energy state. Besides emission of γ rays, a nucleus can decay in two other ways called α and β emission, where the names are a legacy from when they were first identified in the 1890's. While γ radiation is emission of a high energy photon, α decay involves ejecting part of the nucleus, consisting of two protons and two neutrons, which is essentially the nucleus of a Helium atom. A β decay involves emission of an electron (or anti-electron) from the nucleus. These electrons are produced in the weak decay of a neutron ($n \to p + e^- + \bar{\nu}$) or it's inverse ($p \to n + e^+ + \nu$); the latter is not possible for an isolated proton because $m_n > m_p$, but it can occur within a nucleus when binding energy changes make it energetically allowed.

The emitted α and β particles are sometimes called α and β "rays", because that was the original terminology when their nature was still not fully understood. That terminology still has a use because it carries additional information about the origin of the particle. Although they are all electrons, the names cathode ray, β ray, δ ray, and photoelectron each tell you something about how the electron was produced. (You're encouraged to look up the details of δ rays if you are curious.)

We'll use Cobalt-60 in this lab. The complete symbol for this isotope is ${}^{60}_{27}\mathrm{Co}$, where the lower number (27) is the atomic number, Z, i.e., the number of protons, and the upper number (60) is the number of nucleons, protons plus neutrons. The atomic number is redundant with the Co symbol, so it is often just called Co-60. However, explicitly listing the atomic number sometimes helps when looking at decays to remind yourself of how many protons there are. I'll use the various naming options interchangeably.

Cobalt-60 decays to Nickel-60 through the processes illustrated in Fig. 3. This diagram is structured like an energy level diagram, where the *relative* energy of the states is shown by the numbers along the right side, specified in MeV. So, if we define the lower energy Ni-60 state at the bottom as the zero point for our energy scale, then the Co-60 nucleus has 2.823 MeV higher energy. There are two intermediate states at energy levels of 2.505 and 1.332 MeV. The lines and arrows (solid, dashed and wavy) represent decay paths, with the corresponding energy change listed next to the arrow along with the type of decay (β or γ). The dashed arrow is a rare process that we will neglect here, so the decay proceeds as follows.

1. The ${}^{60}_{27}\mathrm{Co}$ nucleus β decays to a ${}^{60}_{28}\mathrm{Ni}$ nucleus. A neutron is converted to a proton so the number of nucleons stays constant at 60, but the number of protons increases by one so it is now a Nickel nucleus. That nucleus is left in an excited state, 2.505 MeV above the ground state.
2. The excited nucleus decays to a lower energy, but not yet ground state by emitting a 1.173 MeV γ.
3. A second decay brings the nucleus down to its ground state by emitting a 1.332 MeV γ.

The timescale for the β decay, which starts the cascade, to happen is quite long; the half-life of Co-60 is 5.27 years. What is meant by "half-life" is the time required for half of the nuclei to decay. The

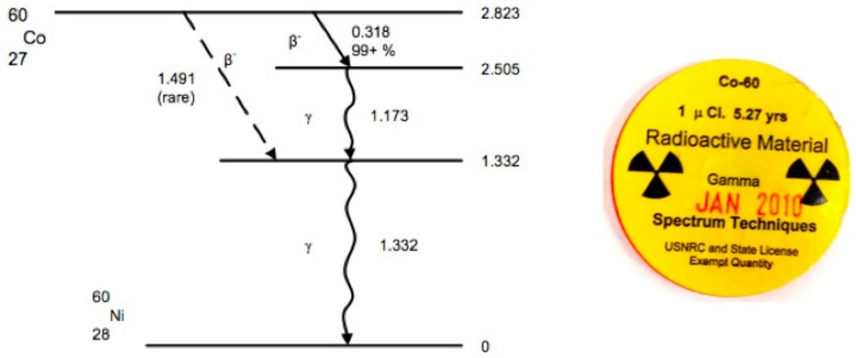


Figure 3: (a) Decay diagram for Co-60. (b) Image of a Co-60 source. The radioactive material is encapsulated at the center of the plastic disk; the yellow label specifies information about the source.

two γ decays happen very quickly afterward, so although they are in fact decays of a Nickel nucleus, we still call this a Co-60 source. If the decays were to happen slowly, we could isolate the material part way through the decay to obtain a different source. That is, in fact, how many radioactive sources are produced; they are metastable products of the decays of other radioactive elements, typically from the Uranium and Thorium decay chains. The Co-60 source that we will use, however, is produced synthetically by bombarding Co-59 with neutrons. It is commonly produced for a variety of applications, including radiation therapy for cancer treatment. The activity of the source is measured in Curie, where $1\mathrm{Ci}$ corresponds to 3.7×10^{10} decays per second. The sources we will use have typical activities of $1\mu\mathrm{Ci}$, or at least they did when they were originally manufactured, but they have since been decaying.

With the appropriate equipment, it is possible to measure the energy of the emitted gammas. This is called gamma ray spectroscopy in analogy to atomic spectroscopy. That won't be today's project. Rather, we will study how the emitted radiation interacts with material. So, I'll begin with a short introduction to that.

2.1.1 Interaction of radiation with matter

Each of the three types of radiation, α, β, and γ, transfer their energy to matter by electromagnetically interacting with the electrons in the material, but they do so with qualitatively different features. Since β particles are high energy electrons, they will scatter off the electrons in the material's atoms, ionizing them in the process, and losing energy. In a similar way, α particles lose energy by ionization, but it happens much more quickly because the α's have charge of $2e$ and have much higher mass than electrons. The rate for energy loss through ionization is proportional to the square of the particle's charge, and inversely proportional to the square of its velocity. As a result, α particles lose all their energy in less than a mm thick layer of most materials, and β's will stop after traversing a few mm. (The rate of energy loss is specified by the Bethe-Bloch formula, the details of which are beyond the scope of this lab).

Gamma rays interact in several different ways, including ionization of atoms via an "atomic photo-electric effect"; Compton scattering off atomic electrons, which ionizes the atom and gives the ejected electron high enough energy to cause further ionization of its own; and pair production, which involves creation of an electron and anti-electron (positron) that have sufficient energy to further ionize. Each of these processes depends differently on energy. The first two processes dominate for the ~ 1 MeV gamma rays from Co-60 that we will use; pair production becomes important only at higher energies. The rate for these interactions is much smaller than for the β interactions, and gamma rays can penetrate several cm of material before interacting. When they do interact, they typically deposit all their energy in a small volume at the interaction point. Because of this feature, we can describe their interactions in terms of an absorption length, i.e., the number of gamma rays remaining after traversing a material of thickness x is $N = N_0 e^{-x/\mu}$.

The dominant effect that sets the absorption length is the density of atoms along the gamma's path, which is related to the mass density of the material. We can account for that, and obtain a definition of absorption length that is applicable to a mix of different materials, by defining a "mass absorption length", which has units of $\mathrm{g/cm^2}$, i.e., density times length, and is given the symbol λ. (Don't confuse this with wavelength. While it may be the same symbol, we're using it for a different purpose.) We can also define a "density thickness", z, measured in $\mathrm{g/cm^2}$ rather than cm, so the number of gammas remaining after passing through that thickness is

$$N = N_0 e^{-z/\lambda} \tag{4}$$

If you have a material, such as aluminum, whose thickness x you have measured, then you can get the *density thickness* from $z = x * \rho$ using the known density, ρ, of aluminum.

In this lab, you'll measure the mass absorption length, λ, for the γ's from Co-60.

2.1.2 Apparatus

The apparatus you'll use is called a *Geiger counter*. It consists of a tube of gas with a metalic cathode around the outside and a thin wire anode situated along the center of the tube. A high voltage applied between them produces a large electric field inside the tube. The gas will be ionized by any incident radiation, due to the processes described above. The released electrons and ionized atoms will be accelerated by the electric field toward the anode and cathode, respectively. Since the anode wire is very thin, the electric field near it is very high, and the electrons accelerated by it obtain enough energy to ionize additional gas atoms. This leads to an avalanche of ionization, resulting in a large and detectable current pulse. Counting the number of such pulses over a period of time measures the rate of incident radiation. By varying the thickness, z, of absorber between the source and the Geiger counter, and measuring the count rate $N(z)$ you can use Equation (4) to determine λ.

You will be counting the number of radiation detection events that occur over some period of time. So, what is the uncertainty on the number that you measure? This idea of counting "events" is common to many experiments and has a standard methodology for determining the uncertainties called "counting statistics". Stated simply[2], the uncertainty on the number of events, N, that you count is $\delta N = \sqrt{N}$. This works for counting anything, radiation events or the number of birds that fly overhead per hour. You might protest that there is no uncertainty in the number that you count. If you saw 15 birds, then there were 15 birds. That is correct, but the reason we are counting these things is to measure the underlying event rate. What the uncertainty represents is the possibility of fluctuations in the number of events that could have happened during the counting time, not how many actually did happen. If you repeated the measurement, counting over many different periods of time, fluctuations would be seen between measurements of the number of events in the different time periods. So, δN represents the range that will contain 68% of those measurements. It accounts for the fact that the time of occurrence of these random events is, well, random.

The Geiger counter that you will use is shown in Fig. 4. On the left is an area where the absorber sheets are stored. On the right is an open area with slots into which you can slide the plastic source holder and the absorber sheets. The Geiger counter is situated above the top of that open area, so absorber

---

[2] This is valid when N is not too small, say $N \geq 5$. For smaller numbers, the more complicated Poisson statistics approach is required, but using $\sqrt{N}$ is fine for all your measurements here.

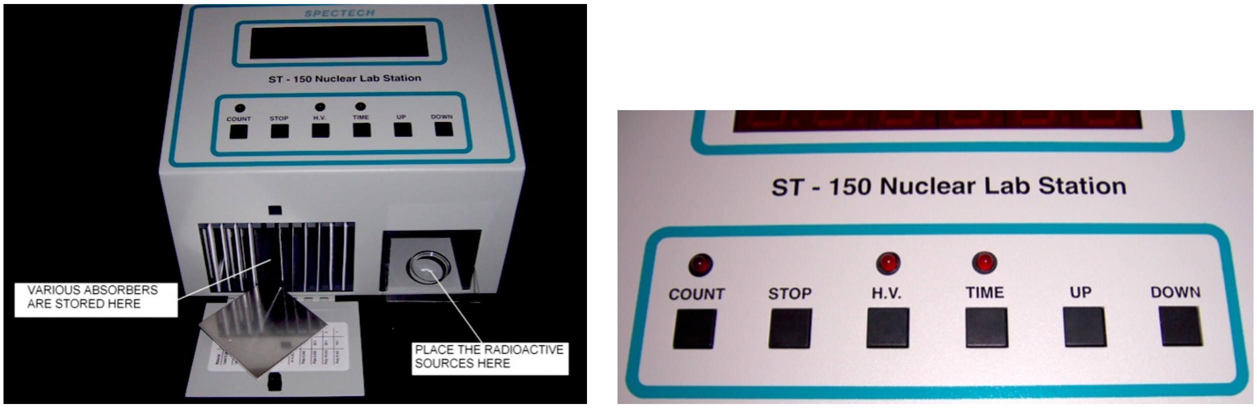


Figure 4: (a) Photo of the Geiger counter apparaturs. (b) Close-up of the control buttons.

sheets placed above the source will be between it and the counter. You can set the length of time for which to count using the *time* button, and then use the *count* button to begin a counting run. It will stop automatically after the selected time period. The number of counts are displayed above the control buttons. Note that when the red LED above the “time” button is lit, the display will show the time rather than the number of counts. To switch back to displaying the number of counts, press the time button again. The “HV” button operates similarly.

2.1.3 Safety and care of equipment

While the radioactive sources that we are using have a low activity, it is appropriate to follow standard safety protocols when dealing with them. These, and other precautions, are listed below.

- *Do not damage the sources.* The radioactive material is encapsulated in the plastic disks. Releasing the material, e.g., by physically breaking the plastic, could cause contamination that is difficult to contain.

- *Do not eat or drink in lab, and wash your hands when you leave the lab.* These protocols are required in any experiment, and their motivation is illustrated by the details in this case. If your hands became contaminated by radioactive material or lead dust, which is also hazardous, you don’t want to let contaminants into your body by eating, either in lab or afterward before you have washed your hands.

- *Do not lose the sources.* The sources are kept in a locked box at the front of the room. You can check one in and out by contacting the TA.

- *Do not touch the active part of the Geiger counter.* There is a thin window up in the top of the open shelf area, which covers the active part of the Geiger Counter. It is kept thin to minimize any absorption of α or β particles. To avoid damaging the window, do not put anything, like your fingers, up into the region above the shelf area.

- *Don’t set the voltage for the Geiger counter too high.* To avoid sparking damage, the operating voltage for the Geiger counter should be limited. They operate best at 500 V; set the HV to that value, but no higher.

2.2 Getting started and gaining familiarity

Before beginning to collect data for the measurement of λ, you should do some simple measurements to see the basic operation. With the HV set to 500 V and the count time set to 60 seconds, measure the count rate per minute with the source placed in the top slot, i.e., as close to the Geiger counter as possible. Do this twice so you can compare the measurements and see the $\sqrt{N}$ fluctuations. Then collect data with the source in two other positions, at the bottom and about halfway up. You should see that the count rate decreases approximately as the inverse square of the distance between the source and the counter.

Next, remove the source, place it at least 50 cm away from the counter, and measure the count rate. That will give you a measure of the *background counts*, i.e., the rate that comes from sources other than your Co-60 source.

For subsequent measurements, you'll want to place the source in a slot toward the bottom to make sure that you have room to stack at least three different absorbers above it. However, keeping it close to the detector allows you to collect data faster. Picking the third or fourth slot from the top seems like a good balance; that allows combining the mass thickness of at least three different absorbers to allow a large range of z.

Those measurements should have given you an idea of the count rate with no absorber. To get an idea of what the count rate will be with absorbers in place, you should take a quick, 60 second long, run with the source in whatever position you've decided upon, and with the thickest set of absorbers available between the source and detector. Now that you know approximately the range of count rates, you should decide on a length for your subsequent runs. You want them to be long enough that the uncertainty on the count rate is less than 5 or 10%, but short enough that you can take many runs with different absorber combinations. The length you choose needn't be exactly calculated, rather it is best to pick something that is about right, and round it to the nearest multiple of 60 so it is easy to calculate counts per minute.

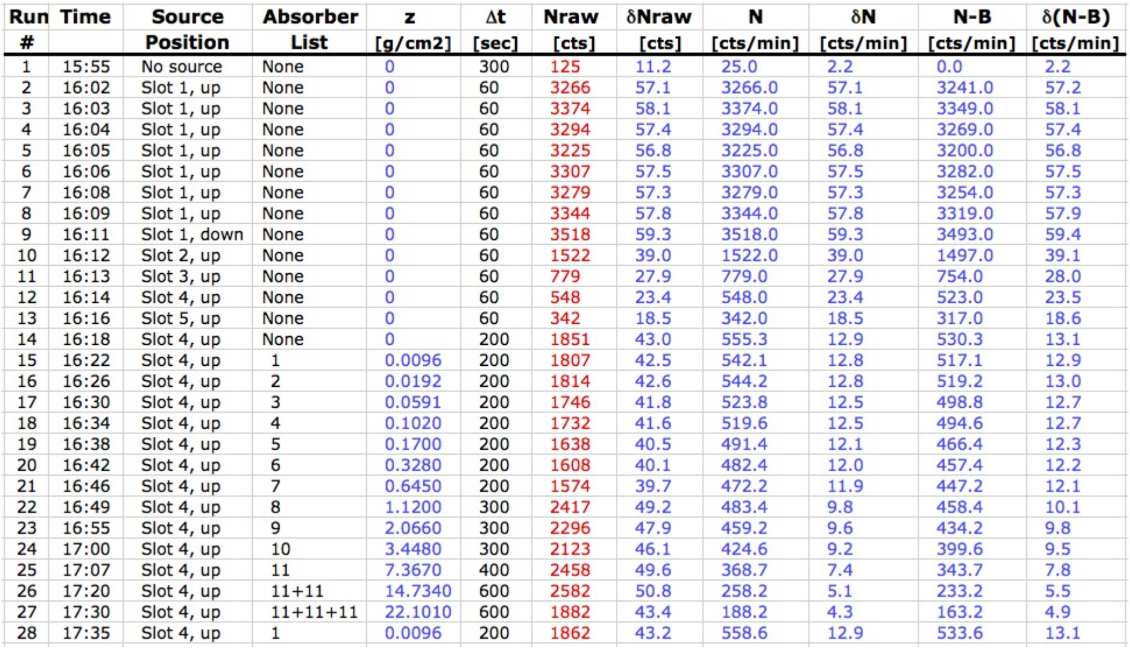
2.3 Measuring the mass absorption length

With the source in your chosen position, collect a set of runs starting with no absorber and then with increasing thickness of absorber. You should measure the count rate with at least 10 different z values, and you'll want to include a large range of z, both very thin absorbers like the aluminum sheets and also very thick ones like a stack of the thickest available. Your apparatus only contains one of each type of absorber, but you can borrow thick lead sheets from a neighbor or the TA; stacking them will allow larger summed z values. The density thickness, z, of each absorber sheet is listed on the door that flips down. The uncertainty on these values is small compared to the effect of the counting uncertainty; you can neglect it.

Make sure not to move the source between runs so that the only thing changing is the absorber thickness.

Record your data carefully in a table. It is a good idea to give each separate run a unique number to help with the bookkeeping. Record the type of absorber(s) used and the total z value. You should record the raw count rate and its uncertainty, then subtract the background counts (measured above) and record the counts per minute as well as its uncertainty in a different column of your table. Since both the raw counts and the background that you subtract have an uncertainty, the uncertainty on the difference between them should be calculated by adding the two uncertainties in quadrature. An example data table is shown in Fig. 5 to illustrate what is expected.

Once you have collected your data, make a plot of the rate, after background subtraction, as a function of z and see if it fits to an exponential function as expected from Equation (4). Make your plot with both a linear scale and with a log scale on the y-axis. A log scale is useful in this case because an exponential



Run #	Time	Source Position	Absorber List	z [g/cm2]	Δt [sec]	Nraw [cts]	δNraw [cts]	N [cts/min]	δN [cts/min]	N-B [cts/min]	δ(N-B) [cts/min]
1	15:55	No source	None	0	300	125	11.2	25.0	2.2	0.0	2.2
2	16:02	Slot 1, up	None	0	60	3266	57.1	3266.0	57.1	3241.0	57.2
3	16:03	Slot 1, up	None	0	60	3374	58.1	3374.0	58.1	3349.0	58.1
4	16:04	Slot 1, up	None	0	60	3294	57.4	3294.0	57.4	3269.0	57.4
5	16:05	Slot 1, up	None	0	60	3225	56.8	3225.0	56.8	3200.0	56.8
6	16:06	Slot 1, up	None	0	60	3307	57.5	3307.0	57.5	3282.0	57.5
7	16:08	Slot 1, up	None	0	60	3279	57.3	3279.0	57.3	3254.0	57.3
8	16:09	Slot 1, up	None	0	60	3344	57.8	3344.0	57.8	3319.0	57.9
9	16:11	Slot 1, down	None	0	60	3518	59.3	3518.0	59.3	3493.0	59.4
10	16:12	Slot 2, up	None	0	60	1522	39.0	1522.0	39.0	1497.0	39.1
11	16:13	Slot 3, up	None	0	60	779	27.9	779.0	27.9	754.0	28.0
12	16:14	Slot 4, up	None	0	60	548	23.4	548.0	23.4	523.0	23.5
13	16:16	Slot 5, up	None	0	60	342	18.5	342.0	18.5	317.0	18.6
14	16:18	Slot 4, up	None	0	200	1851	43.0	555.3	12.9	530.3	13.1
15	16:22	Slot 4, up	1	0.0096	200	1807	42.5	542.1	12.8	517.1	12.9
16	16:26	Slot 4, up	2	0.0192	200	1814	42.6	544.2	12.8	519.2	13.0
17	16:30	Slot 4, up	3	0.0591	200	1746	41.8	523.8	12.5	498.8	12.7
18	16:34	Slot 4, up	4	0.1020	200	1732	41.6	519.6	12.5	494.6	12.7
19	16:38	Slot 4, up	5	0.1700	200	1638	40.5	491.4	12.1	466.4	12.3
20	16:42	Slot 4, up	6	0.3280	200	1608	40.1	482.4	12.0	457.4	12.2
21	16:46	Slot 4, up	7	0.6450	200	1574	39.7	472.2	11.9	447.2	12.1
22	16:49	Slot 4, up	8	1.1200	300	2417	49.2	483.4	9.8	458.4	10.1
23	16:55	Slot 4, up	9	2.0660	300	2296	47.9	459.2	9.6	434.2	9.8
24	17:00	Slot 4, up	10	3.4480	300	2123	46.1	424.6	9.2	399.6	9.5
25	17:07	Slot 4, up	11	7.3670	400	2458	49.6	368.7	7.4	343.7	7.8
26	17:20	Slot 4, up	11+11	14.7340	600	2582	50.8	258.2	5.1	233.2	5.5
27	17:30	Slot 4, up	11+11+11	22.1010	600	1882	43.4	188.2	4.3	163.2	4.9
28	17:35	Slot 4, up	1	0.0096	200	1862	43.2	558.6	12.9	533.6	13.1

Figure 5: An example of a data table to use. This illustrates how you might collect the raw data and calculate what you will later plot. Your numbers will differ, of course. The black entries are set values, the red ones are raw measurements, and the blue ones are calculated values. This table was made with Microsoft Excel; it is useful for tabulating the data and doing simple arithmetic. (If you are not familiar with Excel, you can see how these calculations are done by downloading this spreadsheet from http://stuart.physics.ucsb.edu/l?40515).

should look like a straight line on a log scale. You will likely find that your data does not fit well, which would be revealed by having a large value for χ^2/dof. Inspect the plot to see if you can figure out why. The answer is explained on the next page, but you should be able to figure it out for yourself by looking at the full set of decays in Fig. 3(a) and the discussion of the way β and γ rays interact with matter.

Your plot should show that the data fits well to an exponential function for most of the data points, except for those at the lowest values of z. That occurs because the Co-60 source involves both β and γ decay. The β particles lose energy quickly and can be stopped by a thin layer of absorber. Once they are all absorbed, the remaining γ's follow Equation (4). So, you can best measure λ by excluding the data points at low z from your fit. (If you are using the online plotter, use the *selection criteria* feature to exclude them.) It is a good exercise to try to overlay two graphs (e.g., using the *overlay graphs* feature) to plot the the γ dominated region at high z in one color, with an exponential fit, and overlay the β dominated region at low z in another color. That would show the whole story in a single plot.

2.4 Data analysis

Since this is the first lab, we'll give a detailed example of the type of plots you should make to analyze the data and explain a bit about how. For subsequent labs, the two paragraphs above should be sufficient to guide you, and you'll be expected to develop and execute the analysis on your own.

First, you need to put all your data into the plotter. (See the online tutorial for basic usage.) Here, I'll show plots using the sample data presented in Fig. 5. We'll want to look separately at the measurements with varying z and constant source position, but the plotter allows you to select that subset later, so just put all the data in together. The relevant information is the run number, slot number, z, $N - B$, and $\delta(N - B)$. When entering these into the plotter, you need to give them variable names without special characters, i.e., no minus signs or spaces. I'll call them **run**, **slot**, **z**, **N**, and **dN**, where the subtraction of background is implicit rather than explicit in the names **N** and **dN**.

After entering the data into the plotter, we can plot the count rate, **N**, as a function of **run** for the measurements made with the source positioned label-up in the top slot (runs 2 through 8 in Fig. 5). You can select that data by defining a *selection criteria* of $2 \leq$ **run** ≤ 8 to pick out just the data points you want. The resulting graph of **N** v.s. **run** is shown in Fig. 6(a). Nothing changed between these runs, so they should have count rates that are consistent within uncertainty. Fitting them to a constant (specified by "pol0" or 0th order polynomial) shows what we expect, i.e., about 68% of the data points should have their error bars crossing the best fit line and about 32% should be further away than that. If this were not the case, it would indicate a problem, either because we calculated the uncertainty incorrectly or something was not stable. Another way to validate that the data and uncertainties are appropriate is by looking at the "fit quality" of the reported χ^2 over the number of degrees of freedom (ndf). As explained previously, this ratio should be close to unity. The value of $\chi^2/\text{ndf} = 4.4/6$ is reasonable.

Next, we'll add the one data point taken with the source in the same position but flipped upside down. That is run # 9, which we can pick out using a new selection criteria. Using the "Overlay Graphs" feature, we can plot the data from runs 2 through 8 with a fit as before and then overlay a graph of the count rate for run # 9. That is shown in Fig. 6(b). A different color and marker shape was selected for the run # 9 point to highlight its different configuration. It is useful to take the time to choose a style for your plots that makes them clear with proper axis labeling and choice of colors. Also notice that we turned off the "Stat Option" box that is not useful, and is distracting, in this particular plot.

The run taken with the label side down has a count rate that is higher than the others by about three times the size of its uncertainty. This indicates that there is a significant difference between these two configurations. After completing the lab, you should have enough information to figure out why. A properly skeptical physicist would point out that the difference between run 9 and the other runs could simply be due to inadvertently moving the source to the left or right in its holder while flipping it over. Perhaps that affected the counting rate, regardless of whether it was facing label up or label down. How could you check this possibility? The answer is in this[3] footnote.

Next, we should plot the count rate as a function of z for the data (runs 14-27) with the source held fixed while the absorber thickness was varied. (You could include run 28, though it is mostly useful as a check that there is no significant difference between the start of the data taking period and the end. However, comparing the count rates for runs 15 and 28, we see that they are consistent, so run 28 has served its purpose.)

To make the desired plot, we select the data for these runs with an appropriate criteria and choose a min and max for the plotting range that will display all the data. Figure 7 shows this plot with a linear scale for the abscissa (y-axis) on the left and with a logarithmic scale on the right. To get the logarithmic scale, click the "Log y" checkbox. But beware that you have to change the minimum value for the y-axis to be something greater than zero for any log plots.

In both plots, the data is fit to an exponential function. The resulting fitted function is the same in both cases, but the logarithmic scale in the right hand plot makes it look linear. (That is one benefit of using a log scale.) Notice, however, that the fit quality is poor. You can tell that because the reported χ^2 is quite a bit larger than the number of degrees of freedom, 34.8 vs 12. You can also see that many of the data points are farther away from the fit line than expected based on their error bars. We need to investigate and understand this to proceed with the analysis.

A hint of what is going on comes from noticing that the points at high z match the fit line better than

---

[3] We should have taken several runs where we flipped the source repeatedly to see if the label up and label down runs were consistent with each other despite repositioning and showed a stable discrepancy based only on up/down orientation.

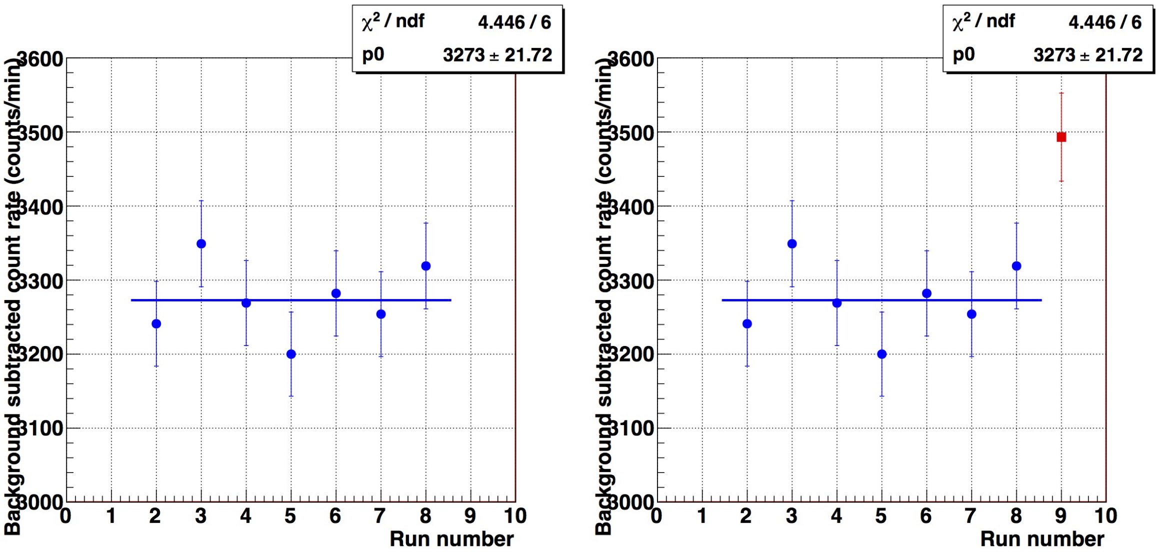


Figure 6: (a) A graph of the count rate vs run number taken with the source positioned in the top slot with no absorber. (b) The same graph where we have added run # 9 that is identical except that the source was turned upside down. It reveals a significantly different count rate is obtained in that configuration.

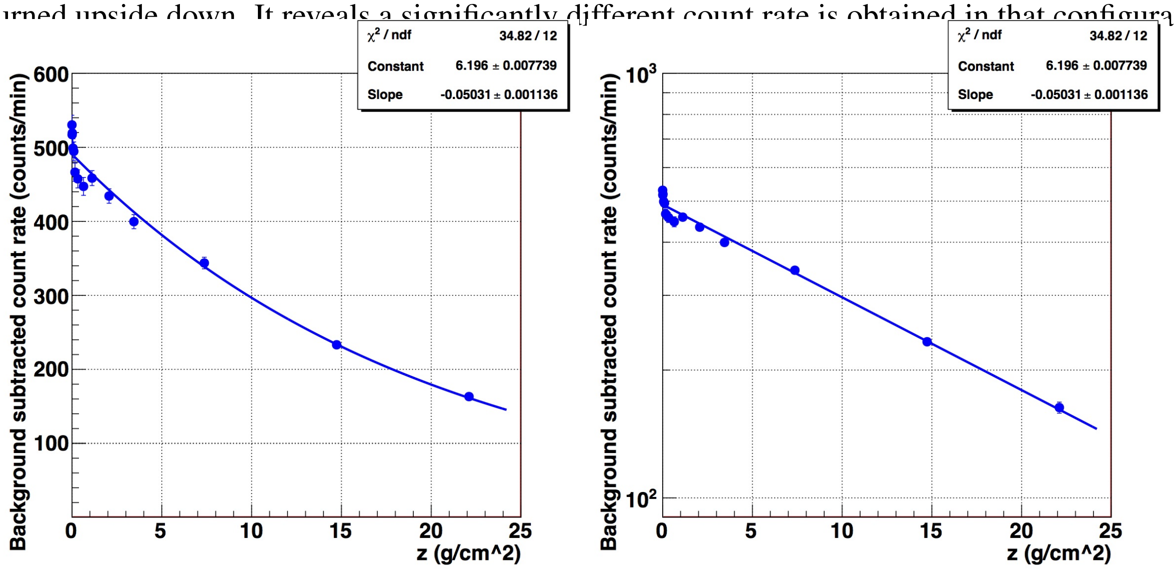


Figure 7: Graph of the measured count rate vs absorber density thickness, z, with an exponential fit, plotted with a linear abscissa scale on the left and a logarithm abscissa scale on the right.

the points at low z. In fact, the largest discrepancies seem to be from runs taken at very low z. Perhaps there is something different about the low z points? Of course, when the fit is done using all the points it will be pulled around by the combined effect of all of them. So we can more cleanly investigate how the low vs high z points behave by fitting them separately. In Fig. 8(a) we use the “overlay graphs” feature to fit the high z data points to an exponential and then overlay the low z points. I decided to separate low and high at $z = 0.5\ \mathrm{g/cm^2}$ after making some test plots where I zoom the x-axis scale into the low z region. Figure 8(b) illustrates this with such a zoomed plot.

These plots show clearly that the data taken with very thin absorbers, z less than about $0.2\ \mathrm{g/cm^2}$, have a different behavior from the data with larger z. The high z data points fit well to an exponential, while the low z points have a different shape that could be a much steeper exponential. You should be able to figure out why. Since we want to measure λ for gamma ray absorption, it is only the high z data that is useful for our purpose. The slope of the exponential fit in Fig. 8 is related to λ. But beware that the fitted function is $y(x) = e^{mx+b}$, where m is the reported “Slope” and b is the reported “Constant”.

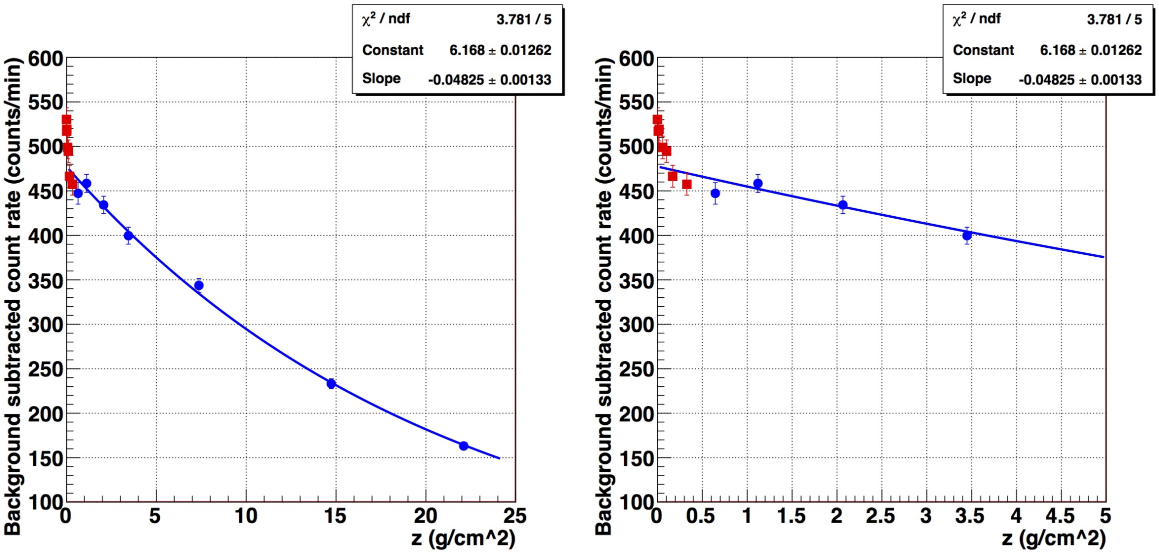


Figure 8: Graph of the measured count rate vs absorber density thickness, z, with an exponential fit to only data points with $z > 0.5\ \mathrm{g/cm}^2$. The left hand plot shows the full range of z, while the right hand plot shows the low z region where a discrepancy is observed for z less than about $0.2\ \mathrm{g/cm}^2$.

So the λ we want from Equation (4) is $\lambda = 1/m$. That is easy to calculate, but remember to properly propagate the uncertainty on the slope to an uncertainty on λ.

2.4.1 Beware biases

Finally, it is worth discussing the points with $0.2 < z < 0.5$ in Fig. 8(b). They are between the two regimes of high z, which we want to use for our measurement, and that of low z that we want to avoid due bias from some other effect. What should be do with them? It is best to leave them out of the fit and use only points well separated from the problematic region, as I have done in the figure. You might note, however, that those points are actually *below* the fit while the bias effect at low z pushes points above the fit. If we were to include those two data points in the fit they would change the slope, so should we include them? In general, you should not. In particular, you should never decide whether to include or exclude a data point based on its value and what it would do to the final result. Such *a posteriori* "cherry-picking" of the data would bias your results!

On the other hand, if our selection of the minimum fit range of $z = 0.5$ had been based on how the data looked around that region, either conciously or unconciously, then we could have already introduced a bias. There are two ways to get around this problem. First of all, don't pick the boundaries too precisely; just choose a round number that looks well separated from the region that you want to exclude. The choice of $z > 0.5$ is a reasonably round number, and so is $z > 1.0$. Secondly, the best thing to do is to collect a new set of data. Once we have figured out from the first round of data that there is a problematic region to exclude, we can choose a boundary to exclude it, either 0.5 or 1.0 would be fine in this case, and collect a whole new set of data to make the final measurement. The fit of this new data will not be affected by fluctuations that might have biased our choices in the first data set. Taking multiple data sets requires more time, but rigor requires effort.

2.4.2 A concise summary of results

When you write up the conclusions for a measurement that you have made, you want it to be concise and yet complete. One way to achieve that is with a single well crafted plot. Figure 8(a) is an example of such a plot. It shows both the low and high z data, clearly separated by color, so it can be used to illustrate the low z bias. It also shows the fit result, which communicates the outcome of the measurement, and the fit quality (χ^2), which illustrates its validity. You should take the time to properly prepare one or two "summary plots" such as this for each of your measurements. Tape printouts into the Conclusions section of your logbook. You should also save these plots with the "Save plot with caption" button. This saving is particularly important for your individual measurement since you'll want to later extract plots electronically to use in the presentation that you will make, and the paper that you will write, about your measurement.

2.5 Going Beyond

There are several ways that exceptional students would go beyond the basics in this lab. I list a few ideas below, but you are encouraged to use your own creativity.

- Check the $\sqrt{N}$ fluctuations by taking repeated measurements and measuring the RMS of their histogram.
- Measure the uncertainty on the thickness of the absorbers. You can do that by measuring the count rate with several different absorbers that are listed as having the same density thickness; use the spread in count rate and the measured λ to measure the variation in thickness.
- Do you see twice as many gammas as betas, as predicted in Fig. 3(a)? If not why not? (Hint the source material is embedded in the plastic disk.)
- Measure the total activity of the source.
- If the source is in the bottom slot, is there any difference between putting a plastic absorber right above it vs in the top slot? Why?
- Carefully investigate whether there is a difference in the count rate when the source is placed label up or label down. Explain this based on what else you observed in the lab. Can you test that explanation somehow?
- Measure the count rate for any object of your choice, like a dollar bill. Is it significantly higher than the background count rate?

3 Interference and Diffraction

Abstract: We will observe the effects of diffraction of visible light through a narrow slit and of interference of light passing through multiple slits. We will use the interference patterns to measure the wavelength of a laser.

3.1 Pre-lab preparation

Interference is a property of waves. Put simply, waves superpose so the amplitude at any position and time is the sum of all waves at that position. If the superimposing waves are periodic and coherent, then they combine in ways that lead to striking phenomena. For example, standing waves were discussed in Phys 23, and diffraction and interference of light waves were discussed in Phys 24. In this lab, we will examine the diffraction and interference of light and use it to measure the wavelength of a laser.

A useful example is to consider two point sources of light that are separated by a distance d and emit completely coherently, i.e., they emit light of the same intensity and wavelength, and do so in phase with each other. That light propagates in all directions, but for simplicity we will consider only a single plane. As sketched in Figure 9, we'll use the (x, y) plane and place the light sources at points $(0, d/2)$ and $(0, -d/2)$. We can then determine the intensity of the light at different points along a screen, placed a distance L away on the x-axis, by considering the superposition of the two light waves at an arbitrary point, (L, y), on the screen.

As shown in Figure 9(a), at the point $(L, d/2)$, the path length for each of the two light waves is equal. That leaves the two waves in phase with each other. So they add (i.e., interfere) constructively, and we'll see a bright spot at that point on the screen.

For another point on the screen, (L, y_{destr}), the path length for each of the two light waves differs as shown in Figure 9(b), where the light that comes from the lower point has an additional path length of $d \sin\theta$. This extra pathlength shifts the phase of the light, by $\Delta\phi = 2\pi d \sin\theta/\lambda$ with respect to the light from the upper point. If the additional path length is $d \sin\theta = \lambda/2$ then that phase shift is $\Delta\phi = \pi$. That makes the wave from the lower point exactly the negative of the wave from the upper point and they destructively interfere. We can calculate the point on the screen where this desctructive interference occurs with, $y_{\text{destr}} = L \tan\theta$. If θ is small, i.e., L is much larger than d, then we can approximate $\tan\theta \approx \sin\theta$, so $y_{\text{destr}} = L \sin\theta = L\lambda/2d$. The point (L, y_{destr}) will be dark.

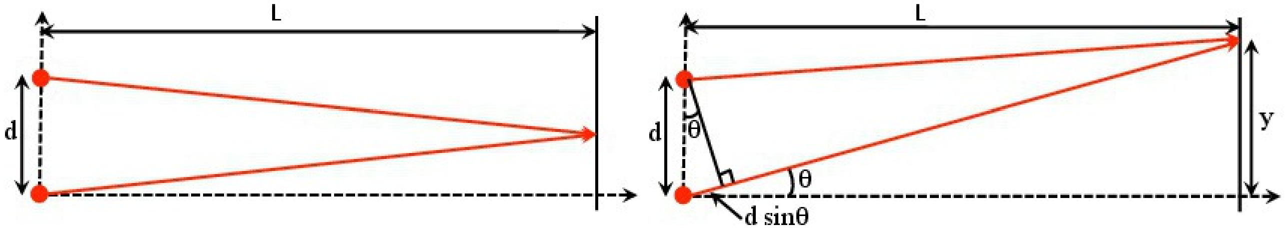


Figure 9: A diagram of two point sources (red dots) used for calculating the interference pattern. The (x, y) coordinate system is shown by the dashed lines. The point sources are the red dots, separated by a disctance d, and the interference pattern is calculated considering the light projected onto a screen a distance L away only the x axis. (a) shows light rays reaching a point of equal pathlength between the two points; here equal pathlength leaves the two waves in phase with each other and constructive interference occurs. (b) shows light rays reach a point where the pathlength differs by an amount $d \sin\theta$, with consequences discussed in the text.

For another point on the screen, (L, y_{constr}), the path length for the two light waves differs by exactly one wavelength, i.e., $d\sin\theta = \lambda$. This shifts the phase by $\Delta\phi = 2\pi d\sin\theta/\lambda = 2\pi$ and makes the two waves equal so they add constructively if $y_{\text{constr}} = L\sin\theta = L\lambda/d$ The point (L, y_{constr}) will be bright.

This repeats for larger values of y. In fact, desctructive interference occurs for any point that corresponds to $\Delta\phi$ being an odd multiple of π, or equivalently $y_{\text{destr}} = nL\lambda/2d$, where n is any odd integer. Similarly, constructive interference occurs at any point where the phase difference is 2π, or equivalently $y_{\text{constr}} = nL\lambda/d$, where n is any integer. Note that $n = 0$ corresponds to the midpoint between the sources that we discussed as the first example. (The drawing in Figure 9 includes a coordinate system offset of $d/2$, but this is negligible since we are working in the limit of $L >> d$.) These relations are key to what we'll do in the lab, so let's rewrite them in an easily findable format:

$$y_{\text{destr}} = nL\lambda/2d, \quad \text{where } n = \pm1, \pm3, \pm5, ... \tag{5}$$

$$y_{\text{constr}} = nL\lambda/d, \quad \text{where } n = 0, \pm1, \pm2, \pm3, ... \tag{6}$$

In fact, you should write these equations, and draw the diagram defining the coordinate system, in your own logbook. You want your logbook to be *a complete standalone description of what you are doing*. Put enough in it so that you can find what you need by flipping pages in the logbook rather than in the lab manual. In fact, spend 5 minutes writing this in your logbook now, before you even go to lab.

An experimentally direct way to arrange for two such coherent light sources is to shine a single laser on something that has two very small holes poked in it. By Huygen's principle, each of these holes will become point sources from which light emanates as spherical waves. And, since the laser light is coherent, the holes will provide two coherent light sources. This is basically the approach that we'll use in this lab, however we'll use *slits* rather than holes. We are only looking in two dimensions, (x, y), and a slit along z is the same as a hole for that purpose. And, we can hit a slit without having to aim as carefully!

A complication in this is that Huygen's principle, which would allow us to treat the holes as sources of spherical waves, only holds for *point* sources. We will make the holes (actually the slits) small, but they are not points. So, let's consider the effect of a finite slit width. We'll do that by considering a single slit of width w as shown in Fig. 10. We'll again consider only the x and y plane, and we can analyze the pattern of interference by conceptually breaking the space within the slit into a series of point sources following Huygen's principle. This is illustrated by the red dots in the figure.

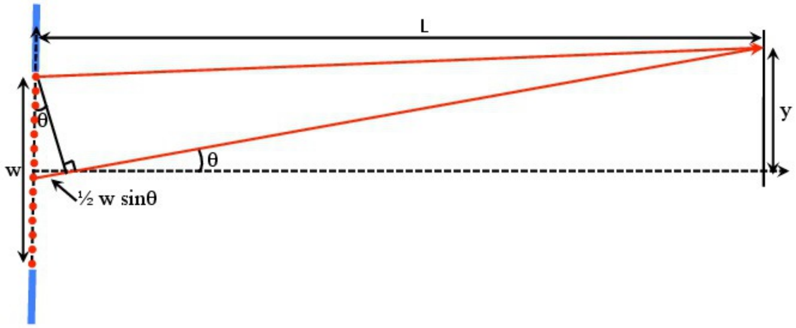


Figure 10: Diagram of diffraction through a thin slit. The blue lines are a material blocking all light except for a thin slit of width w situated symmetrically across the y-axis. We analyze the situation by imagining a large number of point sources within the slit; these are indicated by the red dots. Interference of light from two points separated by $w/2$, one above and one below the y-axis, have pathlength difference of $\frac{w}{2}\sin\theta$ when reaching the screen placed at $x = L$, with consequences discussed in the text.

It is easy to see that there is constructive interference at the point $(L, 0)$ because for each of our imagined point sources in the slit, at a position of say $(0, y_1)$, there is another point $(0, -y_1)$ that is symmetrically across the y-axis and therefore has the same pathlength to the point $(L, 0)$. As a result, there is some constructive interference, and hence a bright spot, at $(L, 0)$.

Using a similar argument, we can see that there is destructive interference at a point (L, y_{destr}). This arises because light from the point at $(0, w/2)$ and $(0, 0)$ have a pathlength difference of $\frac{w}{2}\sin\theta$. If that equals $\lambda/2$, then the light rays combine at $y_{\text{destr}} = L\lambda/w$ out of phase and desctructively interfere. (Note that the two factors of $1/2$ cancel, and we have again used the small angle approximation.) So the light from the point $(0, w/2)$ is canceled by the light from $(0, 0)$, a point $w/2$ below it, and vice versa. And, in fact, the light from any other point in the upper half of the slit is canceled by a corresponding point from the lower half of the slit. That makes the point y_{destr} have completely destructive interference; the point appears dark. The same argument applies for any point that is an odd integer multiple of $\lambda/2$, so there are dark spots for

$$y_{\text{destr}} = n\lambda L/w, \quad \text{where } n = \pm 1, \pm 2, \pm 3, ... \tag{7}$$

This is called a *diffraction pattern*; it is interference of the light passing through different parts of a finite width slit. The light is said to “diffract” through the slit. (Note that as the slit width w goes to zero, the point of destructive interference expands out to infinity and we get the spherical wave pattern of a single point. As w goes to infinity, the spacing between dark points goes to zero so no pattern is discernible.)

In our experiments, we’ll look at both the diffraction pattern from a single slit and the interference pattern from two, or more, slits. The pattern from two slits will be a convolution of the interference from the two slits and the diffraction from each slit. The patterns expected are sketched in Figure 11.

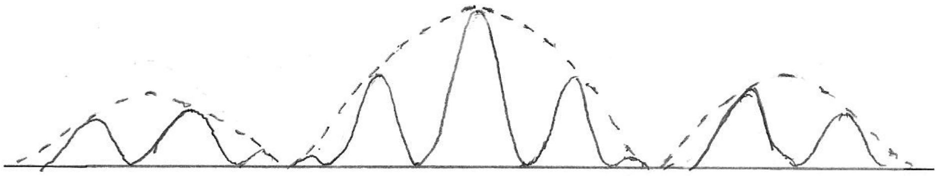


Figure 11: Cartoon of the pattern expected from two thin slits. The dashed line indicates the diffraction pattern, which is convolved with the interference pattern to give the intensity vs position shown by the solid line.

3.1.1 Experimental planning

In this lab you will observe the qualitative features of the interference and diffraction patterns and then use them to measure the wavelength of the laser. Since the wavelength of light is less than a micron, we’ll need to use some experimental tricks to extract a precision measurement. The interference itself provides a large lever arm because the pattern scales with the wavelength times L/d. Using small slit spacings, d, and a long optical path, L, will make the pattern visible. But, it will still only be a few centimeters across. How can we precisely measure the positions of the maxima and minima? Think for a minute or two before reading further to see if you can come up with an idea.

We will also need to measure the distance between the slits. The apparatus provided will specify what the width and spacing is for each set of slits. The range of available widths is between 0.02 mm and 0.16 mm, and the range of spacings is from 0.25 to 0.5 mm. We could just use the advertised sizes, but for a serious measurement you don’t want to just trust what the sales department says! You should

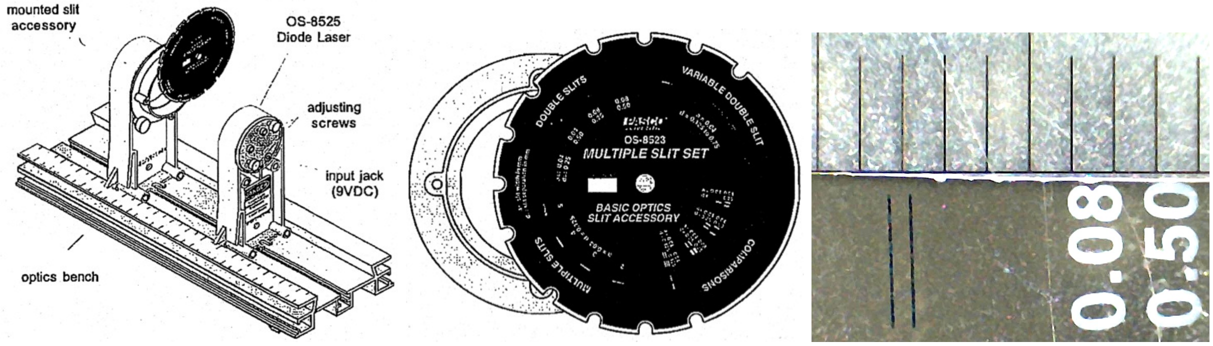


Figure 12: (a) The PASCO OS-8515 Basic Optics System with a OS-8525 Laser Diode assembly and OS-8523 slit accessory mounted. The Multiple Slit Accessory is shown in (b). There is a similar Single Slit Accessory; each of these allow you to change which slit (or set of slits) is in line with the laser beam by rotating part of the device. A closeup view of one pair of slits, taken through a microscope, is shown in (c); a mm scale ruler is placed at the top of the field of view to measure the scale.

measure it yourself. How can you measure these small distances with precision of 10% or better? Again, think a bit about your own ideas for this before reading further.

3.1.2 Apparatus

The apparatus that we will use is an *optics bench* with movable laser, slit and screen mounts, as shown in Fig. 12(a). The "bench" provides a robust way to set and measure the positions of each of the movable mounts. (Attach a piece of graph paper to the metal screen with plastic clips; don't write directly on the metal screen.)

Measuring positions of maxima and minima: One way to measure the positions of the maximu and minima in the interference pattern is to take a photo of the pattern with a ruler placed in the field of view. You can then use image editing software to measure the spacing of the maxima and minima in units of pixels within the image and scale that by the number of pixels per mm using the ruler. In fact, you can use ruled graph paper as the background for the screen and then you don't need a ruler.

Your cell phone is a convenient tool for taking the photos. You can use the computers in the lab room(s) to do the image viewing and pixel measurements, but you may prefer to bring your own laptop and a USB cable to connect your phone for data transfer. If you are using a Mac, then the Preview application has a simple way to select points on an image and determine their relative coordinates in pixel space. However, you may also want to use the Imagemagick suite of tools that allows you to rotate and otherwise modify images. (It is also a way to begin to gain some familiarity with the unix operating system that underlies your Mac. If you are not familiar with that, it is a good idea to begin to learn about this tool that will be handy for future data analysis.) If you are already familiar with a programming language, feel free to use whatever you have in your tool belt.

Measuring the slit width and spacing: You will also need to make a precise measurement of the slit width and spacing. This could also be done with a camera and image processing, *if you have enough zoom* in your camera. But, even then you'd have a hard time holding the camera still enough to get a well focussed image. So, we have a couple USB microscopes that will serve the same purpose. They're basically a webcam with microscope optics attached. They have enough magnification to resolve the slits with a reasonably large number of pixels. So, you can get an image with a length scale (ruler) within the image, and process the image as discussed above. Note that the magnification and resolution is probably insufficient to measure the slit width. However, the spacing can be measured reasonably well, and that is what is important for an interference pattern.

3.1.3 Safety

The lasers that we will use in this lab are not high power, but they can still damage your eye if the beam strikes your eye directly. A few safety precautions are appropriate:

- Do not look directly into the laser beam.
- Use caution when adjusting the optics to avoid directly reflecting the laser toward your, or anyone else's, eyes.

Plan to bring a camera or cell phone and, if possible, a laptop that you can connect to it.

3.2 Getting started and gaining familiarity

The first thing you should do in lab is play around with the optics bench to make sure you understand how it works. Slide the mounts around, and set up your screen. Then take a photo of the setup with your cell phone and tape it into your logbook. You don't need to immediately printout the photo and tape it in. Just block off a spot for it and record the time you took the photo (as a reference to find it on your phone later). Then, before the end of lab, print it out and tape it in.

3.3 Qualitative observations of interference and diffraction

Use the Single Slit Accessory to view the diffraction pattern using at least three different slit widths. Take photos of the patterns and put them in your logbook. Again, you should block off a spot in your logbook when you take the photo and record the time. You can actually print it out and tape it in later. You might want to write your name and the slit width on the graph paper and include that writing in the photograph's field of view. In fact, it is generally a good idea to include such "metadata" directly in your experimental data whenever possible to resolve any ambiguities later.

Write a brief qualitative summary of your observations. For example, you might conclude that the separation of the maxima and minima increases as the slit width decreases. Of course, that is what you expect from equation (7), but the point of doing an experiment is to record what you observe in a clear and concise way. Imagine that the theory of the diffraction pattern where not yet understood, and write your conclusions with the intention of defining the experimental facts that a developing theory must explain.

3.4 Measure wavelength of laser

The observations you just completed are a qualitative way to get familiar with the phenomenon. Now, we get to the measurement. Your goal is to measurement the laser's wavelength as precisely as possible. You'll do this using the the Multiple Slit Accessory and the relations in equations (5) and (6), where you can measure L, y, and d and extract λ. (Recall from section 3.1.1 that you should *measure* d not just take for granted the advertised values.)

You should start by thinking about the uncertainties. In your logbook, calculate the uncertainty on λ in terms of the uncertainties on the things you will measure. What do you think the uncertainties will be on each of the three inputs, i.e., δL, δy, and δd? Which of the three do you expect to dominate the total uncertainty on λ?

The uncertainty on the slit separation, δd, is likely to be a dominant uncertainty. So, it is worth putting some time into measuring it carefully. We have available a couple USB microscopes and some zoom lenses that you can fit over your cell phone camera. You'll want to take several different photos and obtain multiple measurements. Averaging these measurements gives you a more precise result, and looking at the spread of the measurements helps you determine a valid uncertainty. You can make the slit measurement at the beginning or end of the interference pattern measurements, however you prefer.

Next, you should think about systematic uncertainties. While the fractional uncertainty on L is probably going to be smaller than the other uncertainties, it is something that could easily have a systematic bias. For example, while the optics bench provides a nice way to measure the position of the little tab that sticks off the side of the mounts, that isn't necessarily the position of the slit or the screen. The difference between those two positions will be a systematic bias on L, and how precisely we can determine the size of that difference is a systematic uncertainty on L.

There is, however, a trick to avoid that systematic uncertainty on L. We should measure y as a function of L and fit it to a straight line. The slope of that line gives us λ (after correcting for d). The systematic uncertainty from the difference between the mount's position and the actual slit position cancels out in the fit–well actually, it just causes the line to have a non-zero intercept. Measuring the dependence of two observables and extracting a desired quantity from a fit is a common trick because of this advantage and because it is a very good way to validate the data. For example, the goodness of the fit (the χ^2 per degrees of freedom) shows whether the measurements are in fact behaving properly within the determined uncertainties. If the χ^2 is too small, then the uncertainties may be overestimated. If the χ^2 is too big, then either the uncertainties are underestimated or there is something else going on beyond what is assumed in the fit.

You should make measurements for several different values of L and use a fit to extract λ.

As discussed in section 3.1.2, you can obtain a more precise measurement of the positions of the maxima and minima by photographing the pattern and using image editing software to determine the positions in pixels that are then scaled to distance. Think a bit about how you will do this, i.e., plan your experimental protocol before you begin. Examples of good procedures are:

- Make a table in your logbook where you record information about each measurement. It should include a unique measurement or photo number.
- Write that number on the graph paper (attached to the screen) so it is visible in the photos as metadata.
- Note the filename of each photo in your logbook. (You needn't print out and tape in these photos, but save the files somewhere you can find at some far future time when you might need to revisit this.)
- Take more than one photo for each value of L. This will allow you to select the best one, e.g., if some have focus problems. It will also allow you to compare the measured y values between two or more good photos to assess whether the uncertainties are as you estimate.
- It is more precise to measure the difference between the +1 and -1 maxima and then divide by two than it is to just measure the difference between +1 and 0. Similarly, you might find it more precise to determine the position of the minimas than the maximas.
- Leave enough space in your table to add comments or additional information later.

- Using a larger number of slits gives narrower, more sharply defined maxima.

Once you have thought through your plan, write a brief synopsis of what you decided and why into your logbook and then start taking data.

Ideally you would analyze each photo as you take it, but that may be difficult. You should at least promptly analyze the first photo, i.e., copy it into your image processing software and extract the measurements to make sure that everything works well. It is best to find problems at the beginning rather than after you've spent 30 minutes collecting data.

3.4.1 Image analysis

You are free to choose whatever method you prefer to measure the positions in your images, but here are some potentially helpful hints.

The Imagemagick software suite allows you to rotate and otherwise modify images. While you can download it for both Windows and Mac operating systems, a Mac (or linux) operating system is assumed below where it matters for the details. A few handy commands, which you can enter into the Terminal application, are below:

-
  ```
  display File1.jpg
  ```
  This will open a program that allows you to modify and save the image. The program uses the X11 graphics system that is a part of unix. You may not have it installed, in which case you can instead use the convert tool discussed below.

-
  ```
  convert -rotate 5 File1.jpg tmp.jpg
  ```
  This will rotate the image by 5 degrees. It is useful to align the thing you want to measure with the horizontal or vertical axis. You can either calculate the necessary rotation from measurements of the original image or just use trial and error. To avoid wasting time clicking around in the Finder, you can use the command line to do this. For example, if you placed your images in a folder on your desktop called Phys25L, then the following commands would rotate and view a particular image.
  ```
  cd ~/Desktop/Phys25L
  convert -rotate 5 File1.jpg
  open tmp.jpg
  ```
  Once you are happy with the rotated image, you can rename it to something that you will save, e.g., with
  ```
  mv tmp.jpg File1rotated.jpg
  ```

- If you want to see a summary of available commands, simply type convert without any parameters.

- It may help to enhance the color variation of the image. You can do this with
  ```
  convert -gamma 1.2 File1.jpg tmp.jpg
  ```
  or
  ```
  convert -gamma 0.6 File1.jpg tmp.jpg
  ```

-
  ```
  convert File1.jpg tmp.txt
  ```
  will produce a text file containing the x,y coordinate of each pixel and its intensity in all three colors (red, green and blue). You might find it useful to read through this file to find maxima and minima. Besides looking through it in a text editor, you can use the tools built into unix to extract the information you want. For example, try this:

```
cat tmp.txt | grep -v ImageMagick | awk -F’)’ ’{print $1}’ | awk -F’(’ ’{print $1$2}’ |
awk -F’:’ ’{print $1”,”$2}’ | awk -F’,’ ’{print $1” ”$2$3}’ | sed ’s___’g | sed ’s___’g |
sed ’s___’g > tmp.dat
```

to get a distilled set of data.

Once you have the image set for measurement, you can use the Preview application and the rectangular selection to measure the pixel locations. Althernatively, you could drop the image into a power point slide and draw rectangles on top of it that hit the points you want to measure. From the size of the rectangle, you extract a measure of the size of the object.

Feel free to brainstorm with your classmates how to best get at the information you want. If you get a particularly good idea from a classmate, feel free to use it for your measurement, but you should cite them in your logbook as the source of the idea.

3.4.2 Putting it all in your logbook

Whatever analysis method you use, the goal is to obtain a set of measured y_{destr} or y_{constr} values. These should go into a table in your logbook. And you should make plots of them (e.g., y_{destr} vs L with a linear fit). **Print out the plots and put them into your logbook.**

Finally, make sure that your logbook has a clear and concise conclusion. It should state the qualitative features you observed and the wavelength value that you measured, *including the uncertainty* on your measurement.

3.5 Going beyond

There are several opportunities for you to go beyond the minimal requirements, and I encourage you to do so. This class will prepare you for an experimental research position or internship and ultimately for experimental research in graduate school. I’m often asked by students of lab courses to write letters of recommendation for their graduate school applications. Saying that they did everything required of them is underwhelming, but it is great when I can point out where they have pursued additional ideas and/or tried their own methods. You should make that your habit.

Here are a few additional things that you could measure or explore if time allows.

- Measure the positions of the minima in the photos of your single slit diffraction patterns and plot them vs the slit width. Does it fit to $1/w$?
- Measure the spacing between the first, the second, and (maybe) the third minima in either diffraction or interference patterns. Are they equal?
- Observe the patterns from the other features on the Single Slit Accessory, e.g., there is an opaque line that can be compared to a transparent slit.
- A CD or DVD consists of circular tracks with a small spacing. Light passing through (or reflecting from) those tracks will generate an interference pattern. Now that you know the wavelength of your laser, determine the track spacing for a CD or DVD.
- Measure the thickness of your hair.
- Extract the intensity vs position from one of your interference patterns and use it to make a histogram of $I(y)$.

4 The Photoelectric Effect, Photons and Planck's constant

Abstract: Two critical experimental observations led to the development of Quantum Mechanics, the blackbody spectrum and the photoelectric effect. In this lab we will explore the photoelectric effect and use it to measure Planck's constant.

4.1 Pre-lab preparation

Quantum Mechanics was developed in the early 1900's, but the evidence for it began to appear even in the late 1800's. For example, Heinrich Hertz noticed that sparks started more rapidly when the electrodes were illuminated with ultraviolet (UV) light. Other investigators later learned that free charges, i.e., electrons, were emitted from metal electrodes illuminated by UV light. These are called "photoelectrons". Some metals, mostly the alkali metals in the first column of the periodic table, produced a current of photoelectrons (called a "photocurrent") when illuminated by either UV light or visible light, but other metals only did so for UV light. Detailed measurements of this "photoelectric effect" found the following characteristics:

- For a given metal, there exists a definite cut-off frequency ν_0. Incident light with frequency below ν_0 produces no photocurrent regardless of its intensity.
- For light with frequency, $\nu > \nu_0$, the magnitude of the photocurrent is directly proportional to the intensity of the incident light.
- The maximum kinetic energy of the ejected electrons does not depend on the intensity of the light, but it is proportional to the frequency of the incident light.

These characteristics could not be explained by the classical description of light as a wave which continuously transmits energy that is proportional to the intensity of the light. Explaining these observations required a different description for light, which was proposed in papers by Max Planck (1901) and Albert Einstein (1905).[4] The new idea was to consider light to consist of discrete bundles of energy, rather than the previous view of it being a continuous stream of energy. These bundles, or "quanta", of energy are called photons. Each photon carries an energy of $E = h\nu$, where h is Planck's constant with units of energy per time. The name choice arises because Planck first used this quanta idea to explain the blackbody radiation spectrum. Einstein then used it to explain the photoelectric effect as follows.

A photon in the incident light transfers all of its energy to an electron in the metal. If that energy is sufficient, it can dislodge the electron from the metal, i.e., it does some amount of work, W, to overcome the potential energy binding the electron in the metal. Any excess energy goes into the kinetic energy of the ejected electron, i.e., $E = W + \frac{1}{2}mv^2$. Using the photon energy, $E = h\nu$, we get $\frac{1}{2}mv^2 = h\nu - W$.

Not every electron is in the same potential, so the work required to eject them varies and so their resulting kinetic energy varies. However, there is a minimum required work for the most loosely bound electrons, which corresponds to a maximum kinetic energy. We can measure that maximum kinetic energy by applying a "stopping voltage" that the released electrons must overcome to be detected. If we call V_s the voltage required to stop the most energetic electrons, and then of course all the other electrons as well, then $eV_s = \frac{1}{2}mv^2_{\mathrm{max}}$. Putting this all together we get the expression

$$V_s = h\nu/e - \phi \tag{8}$$

---

[4] The quantum explanation for the photoelectric effect is described concisely in a three page section of Einstein's paper, Ann. Physik 17, 132 (1905), and was the reason for his subsequently being awarded the Nobel Prize.

where ϕ is called the "work function" of the metal and corresponds to the minumum work required to eject an electron from that particular type of metal. Since we divided through by e, the work function ϕ has units of volts.

Equation (8) is a straight line with a slope of h/e and a y-intercept of $-\phi$. By measuring V_s for different frequencies of light and fitting such a line we can determine both Planck's constant and the metal's work function.

4.1.1 Apparatus

The apparatus that we will use is shown in Fig. 13. It consists of a mercury discharge lamp that is the light source; a set of filters that can each select a narrow range of wavelengths from that light; and a vacuum phototube containing a cathode, from which a photocurrent will be produced, and an anode that will collect and measure the photocurrent, as well as provide an adjustable stopping voltage to impeded the current. Using different filters allows you to change the ν of the light. For each filter you can then adjust the anode-cathode voltage and find the value, V_s, that stops all photoelectrons from reaching the anode. That requires measuring the photocurrent that flows between the cathode and anode in the vacuum tube; you can do that with the "electrometer", which is a sensitive ammeter that can measure down to the picoampere range (1 pA = 10^{-12} A). A voltmeter is also provided to measure the impeding voltage applied between the anode and cathode. A diagram of the circuit you'll use for that is shown in Fig. 14, along with a more detailed view of the phototube.

Light enters the phototube and strikes the cathode, which is a metal plate coated with a Cesium-Antimony alloy. That alloy is chosen so that photoelectrons can be ejected from it by a range of visible light wavelengths without having to rely on UV light, which involves hazards worth avoiding. A small fraction of the photoelectrons ejected from the cathode will be directed toward and hit the anode. They'll generate a photocurrent that runs through the electrometer back to the cathode. (This is converting the photons into an electric current like a solar cell does, but it is very inefficient.) The battery and adjustable voltage divider in Fig. 14(a) apply a voltage between the anode and cathode that impedes that natural flow of the photocurrent. It pushes the electrons back from the anode to the cathode, so only electrons with sufficient kinetic energy to overcome the eV potential from the battery can reach the anode and flow through the electrometer back to the cathode.

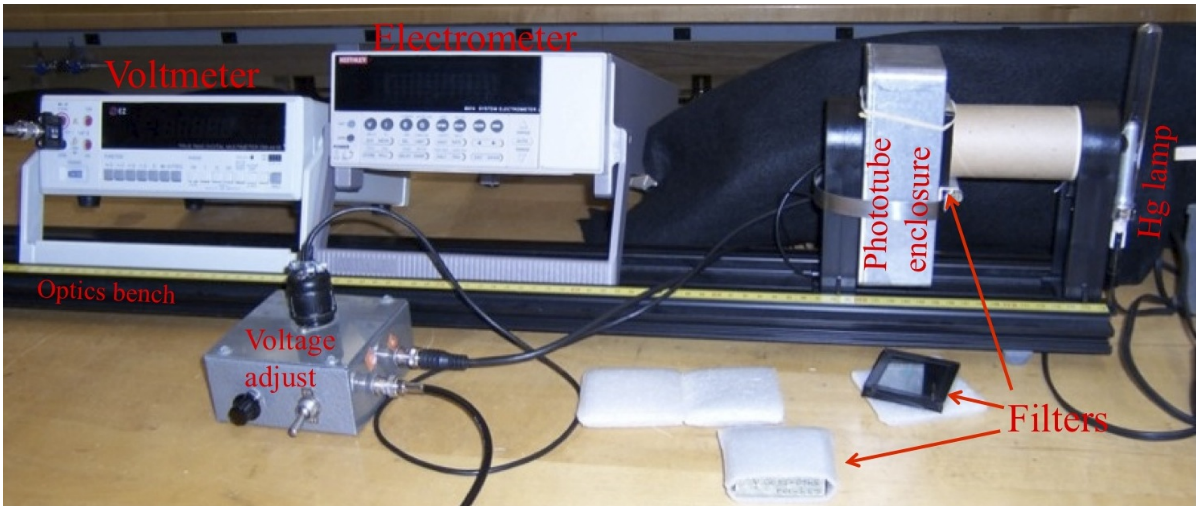


Figure 13: Photo of the apparatus mounted on an optics bench.

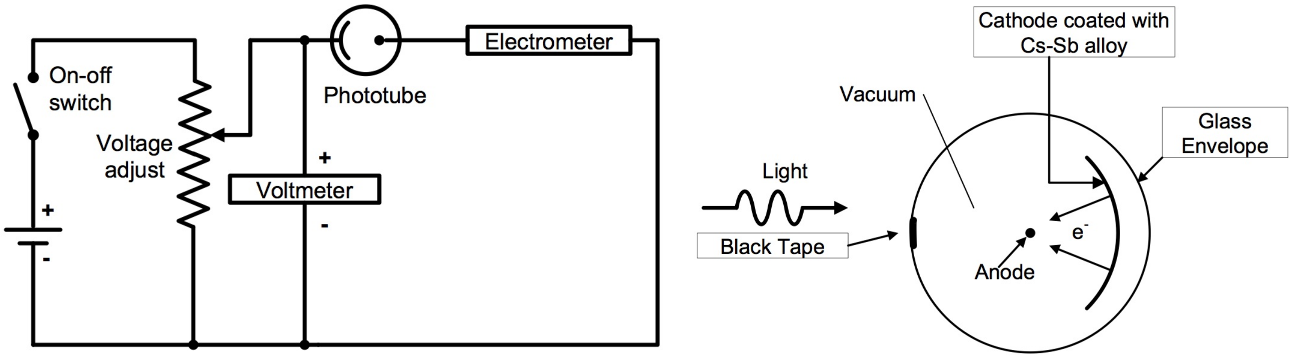


Figure 14: (a) A diagram of the circuit used in the experiment, showing the adjustable voltage divider, which controls the impeding voltage, and the connection of the voltmeter and electrometer. (b) A diagram showing the phototube contents and operation.

Note that Fig. 14(b) shows that a piece of black tape is situated in the light path so as to keep the anode in its shadow. Think about why that is necessary, and write a brief explanation in your logbook.

The phototube is evacuated. Think about why it is important to have a vacuum between the cathode and anode, and write a brief explanation in your logbook.

4.1.2 Safety

The Hg lamp presents two hazards.

- It generates a spectrum of light including emissions in the ultraviolet that are hazardous to your eyes and skin with prolonged exposure. A test tube placed over the lamp is a simple way to shield the ultraviolet component of the lamp's emissions because most glass is opaque at ultraviolet wavelengths. *Before turning on the Hg lamp, make sure that the test tube shield is in place.* Note also that the test tube has a white marking on it. This is normally used for labeling the contents. In our application is it not useful and indeed could be problematic if it blocks the light path. Keep the test tube rotated so that this white mark is out of the way for your light path.
- It gets hot. *Don't touch the lamp or test tube after it has been turned on, and keep any other materials away from contact with it.*

You'll want to make sure that the lamp is properly situated and turn it on at the beginning of the lab. Then leave it on and untouched. That way it will warm up and its output will remain more stable.

The phototube is enclosed in an aluminum box that has a small window for the incident light. This prevents ambient light, which has not been carefully wavelength selected by the filter, from affecting the measured photocurrent. Similarly, a short cardboard tube is used between the filter and the light source. It keeps your hands away from the hot lamp while swapping the filters and masks the ambient light. The battery and adjustable voltage divider are placed in another aluminum box with connectors for the voltmeter, electrometer and phototube. You won't need to open either of these boxes, and you shouldn't. To satisfy your curiousity, however, photos of their contents are shown in Figure 15.

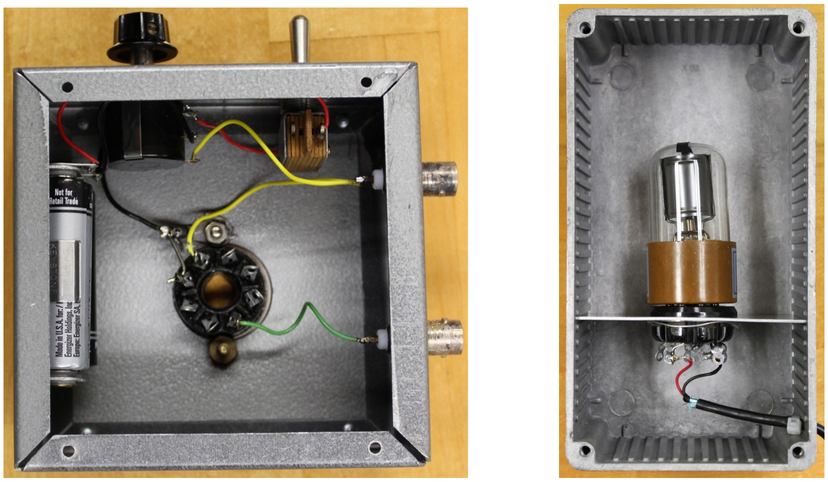


Figure 15: Photos of the aluminum boxes containing (a) the voltage adjustment and (b) the phototube.

4.2 Getting started and gaining familiarity

The first thing you should do in lab is turn on the meters so that they have time to warm up. Most equipment has a “warm up” period, which is simply the time it takes for all the heat production and dissipation mechanisms to come to equilibrium so that the device’s components remain at a constant temperature. Variations in temperature can alter operational details of the electronics, such as amplifier gains and leakage currents. The electrometer has calibration circuitry that measures and corrects for most of these temperature dependent effects, but operating at a stable temperature still improves its precision. Then turn on the mercury lamp–after rereading the safety section. Use the black cloth to make a tent over the lamp and metal enclosure, being sure that it is not touching the hot lamp.

Before taking data, spend some time making sure that you understand how the equipment works and document its configuration. You should take a photo of your setup and tape it into your logbook. Don’t use the photo in Fig. 13 because it is not *your* setup. Remember that the idea of logging photos is to record views of *your equipment* so you can later extract any details about it that might be needed. You should also put a diagram of the circuit setup in your logbook. Here you can use the diagram in Fig. 14, once you have traced the connections in your setup to make sure that they match that diagram.

4.2.1 Operation of the electrometer

The electrometer (Fig. 16) is central to the measurement, and its operation warrants a short explanation. You can find full details by googling for the equipment manual; what follows is excerpted from it.

The Keithley 6514 electrometer can be used to measure voltage, current, resistance or charge. We will only use the current measurement capability. When powered on, it defaults to voltage measurement mode. You can switch to current mode using the front panel button, but you should put the device into Zero Check mode (see below) when switching modes. It has ten ranges for current measurement, with the full scale values varying 20 pA to 20 mA. You’ll need to select the appropriate range for each of your measurements, or use the autoscale feature. For our measurements, the nA scale is probably optimal.

The electrometer is capable of precision measurements of quite small currents, but that precision requires some careful procedures. In particular, you should calibrate and correct for any offset currents. This can be done with the Zero Check and Zero Correct functions. When Zero Check is enabled, the input signal is shorted so that the input amplifier sees only offset currents and voltages that are not from

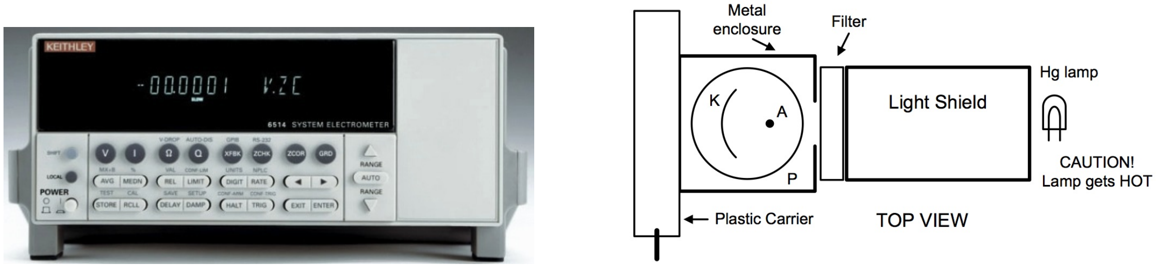


Figure 16: (a) Photo of the Keithley 6514 Electrometer. (b) Diagram showing the position of the filter that will be installed between the lamp and the phototube assembly.

the input signal and thus should be subtracted. That subtraction can, and should, be done internally using the Zero Correct function, following the procedure listed in Fig. 17. It is reasonable to perform the zero correction procedure between each set of measurements, e.g., for each new filter, to compensate for any drifts that may occur over time.

You should also set the RATE setting to SLOW and the DIGIT setting to display 2 or 3 digits after the decimal point. Displaying more digits could provide more precision, but the temporal variation of the measurements may be larger than that and is likely to be what will determine the uncertainty on your current measurements.

Zero Correction procedure to remove offsets

- Enable Zero Check mode. Press ZCHK; you should see a “ZC” message appear at the right side of the display.
- Select the current measuring mode by pressing I (if not already in that mode).
- Select the appropriate measurement range.
- Perform a zero correction by pressing ZCOR; you should see a “ZZ” message appear in the display.
- Press ZCHK to disable the Zero Check mode; you should now see a “CZ” message appear in the display to indicate that the subsequent readings are corrected for any zero point offsets.

Figure 17: Zero correction procedure.

4.3 Measure $V_s(\nu)$ and extract h

There are four optical filters which will pass only light having a narrow range of frequency. By placing these different filters between the Hg lamp and the phototube, you can make measurements of the phototube’s response as a function of the frequency, ν. A filter can be mounted on the small shelf that is attached to the phototube’s mounting box and held in place by a rubber band, as shown in Fig. 13. The filters vary in thickness, so you will have to adjust the space between the cardboard light shield and metal enclosure so that the light shield is touching the filter. Make this adjustment by sliding *only* the plastic carrier holding the metal enclosure; do not move the carrier holding the lamp. After mounting each filter, cover the lamp and phototube box with a tent made from the black cloth to minimize the effects of ambient light.

You can measure V_s for a given filter's frequency by adjusting the impeding voltage until the photocurrent reaches zero; the impeding voltage that stops all the current is the stopping voltage V_s. To obtain a reliable measure of V_s, you should measure the photocurrent, I, as a function of the impeding voltage, V, and fit the data to obtain V_s. Taking a quick, coarse scan of $I(V)$ will give you an idea of the overall shape. You can then decide what voltage steps seem appropriate and record a table of measurements in your logbook. You should then plot each of the $I(V)$ curves and fit them for V_s. Make sure to put these plots in your logbook, including the fit, error bars on the points, and proper labels on the axes.

After you obtain the V_s for each value of ν, fit that to Equation (8) to obtain Planck's constant. This plot is the key result and should certainly be included in your logbook. Note that the fit will also determine the metal's workfunction, which is also an interesting, though less fundamental, parameter.

Although the lab period does not have sufficient time for a thorough exploration of systematic uncertainties in the measurement, you should spend a few minutes thinking about what possible effects might have systematically biased your measurement and what auxiliary measurement you might make to measure these effects, if time allowed it.

You can compare your measured value to the currently most precise measurement, from NIST, of $h = (6.62606957 \pm 0.00000029) \times 10^{-34}$ J s. However, don't let your thought process be biased by how close or far your measurement is from that value. *You should approach your measurements as if you were the first to have ever measured this value!*

Finally, write your conclusion. It should include a statement of what you did, how you did it, and what you measured. Make sure to include the uncertainties on your measurements in the conclusion, and throughout the logbook.

4.4 Going Beyond

There are several ways that expectional students could improve the precision and robustness of this measurement. I list a few ideas below, but you are encouraged to use your own creativity.

- Repeat the V_s measurements with at least some of the filters; if the results differ significantly from the first set of measurements you have evidence for a systematic uncertainty, and you have a measurement of it. That should be included in your final determination of h.

- It would be best to electronically record the currents and voltages to obtain a fine grained $I(V)$ curve. The electrometer has a computer interface that would allow this, but using it requires a lot of setup. Can you think of another, simpler way to record both I and V as fine grained raw data that you could subsequently extract for an $I(V)$ plot with many data points?

- Talk to your classmates and collect their measured values. With all of that data, you could do a global analysis of $V_s(\nu)$ both to understand any systematic uncertainties and to obtain a more precise determination of h.

- We had to rely on the advertised wavelength range for the filters. Can you think of a way to measure that yourself? Even if you don't have time to carry it out in lab, it is interesting to think through and write out a plan for how you might be able to do so.

5 Atomic spectra and the Rydberg constant

Abstract: In this lab we will observe the visible spectrum of Hydrogen, use it to measure the Rydberg constant, and use the spectrum of an unknown source to determine its atomic content.

5.1 Pre-lab preparation

While the quantum explanation for atomic structure was developed in the early 1900's, the evidence for the quantized energy levels of atoms was observed in the late 1800's through measurements of the atomic spectra. In 1885 Johann Balmer measured the visible spectrum of Hydrogen and found that it could be described by

$$1/\lambda = R(1/2^2 - 1/n^2) \tag{9}$$

The ultraviolet and infrared portions of the spectrum were later measured (by Lyman, Paschen, Brackett and Pfund) and found to fit into similar formulas but with integers other than 2 in the first term. They all measured the same coefficient, R, called the Rydberg constant. This was eventually understood within the Bohr model of the atom to be due to photons, of energy $E = hc/\lambda$, being emitted by atoms transitioning between energy states that took on quantized levels of $E_n = -hcR/n^2$, where n is a positive integer. (These energy levels are illustrated in Fig. 18.) That allowed the Rydberg constant to be calculated as $R = me^4/8\epsilon_0 h^3 c$, while previously it had been only an empirically determined value.[5]

In this lab, we'll measure the visible spectrum of Hydrogen as Balmer originally did and obtain a measurement of the Rydberg constant.

Equation (9) applies only to hydrogen. Other atoms have a similar $1/n^2$ structure but with a different coefficient. Each element's spectrum is uniquely characteristic, so spectra are often used to identify the elemental composition of a sample. You'll use this to identify a sample from its spectrum.

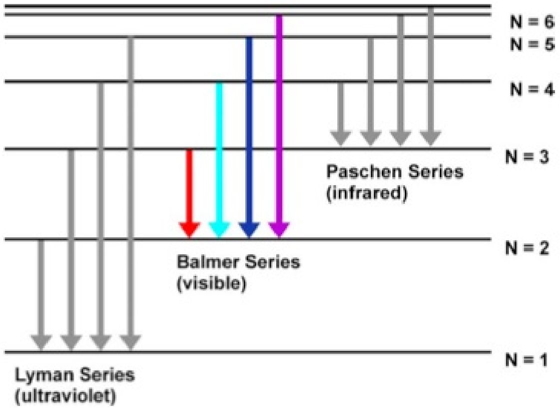


Figure 18: Diagram of the energy levels for a Hydrogen atom. Transitions between excited states down to the $n = 2$ state produce visible light (red to violet, left to right for the four lines shown) that you will observe in this lab. Transitions to the $n = 1$ state produce ultraviolet light, and those to the $n = 3$ state produce infrared light, neither of which can be seen by eye.

---

[5] It is interesting to note that the measured value for R did not match the simplest Bohr model calculation which uses $m = m_e$ because the electron and proton together lead to an effective "reduced mass", m, that is very slightly below the electron mass. Furthermore, spectra of hydrogen showed evidence of a small contribution with a slightly different R value. This comes from deuterium, i.e., hydrogen with a neutron in the nucleus, whose mass alters m. Both effects are smaller than 0.1% and only visible with exquisitely precise measurements.

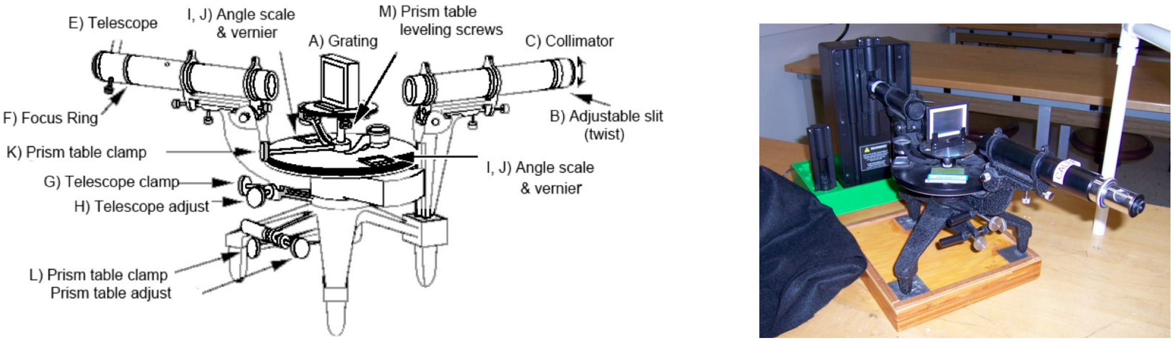


Figure 19: (a) Diagram of the spectroscope with its various components labelled. (b) Photo of the spectroscope. The light source is contained in the square black device behind it.

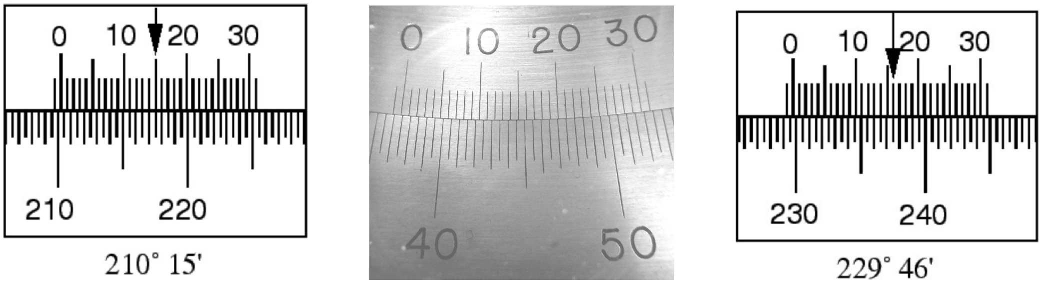
5.1.1 Apparatus

The apparatus you'll use is called a *spectroscope*. It uses interference just like the double slit experiment that you did at the start of the quarter, but rather than two slits we'll use a *diffraction grating* with a very large number of slits. As you should have seen in the first lab, having five slits spaced by a distance d gives a pattern with similar spacing as two slits of the same spacing, however the maxima are much more narrowly peaked, i.e., the contrast is much improved, by the larger number of slits. A diffraction grating takes advantage of this by using a very large number of slits to provide narrow and easily resolved maxima. The diffraction grating that you will use is made by etching thin parallel lines on a piece of glass. The etching process allows hundreds or thousands of lines per mm, which provides a large lever arm on the observed maxima to measure each of the different λ's within the spectrum of the light. Your diffraction grating should specify the number of lines per mm, from which you can calculate the spacing, d, between any pair of lines.

The interference from the slits gives maxima at angles θ for which $d \sin\theta = m\lambda$, where m is any positive integer. The $m = 0$ case is the central maximum with no deflection, i.e., $\theta = 0$. By finding the θ values for the $m = \pm 1$ case, for each of the different wavelengths, you can precisely determine the wavelengths of the emitted light. (You may only be able to see the $m = \pm 1$ maxima. For larger values of $|m|$, the intensity is reduced.)

The spectroscope allows precise measurement of these angles. As shown in Fig. 19, it consists of two cylinders that define the light path. One is a collimator and the other a telescope. The collimator has a narrow opening at its far end that only allows light through a vertical slit, whose width is narrow but adjustable. That slit makes the light beam take the shape of a vertical line. Having the light spread vertically doesn't degrade the resolution on θ, which is the angle in the horizontal plane. Indeed, it is visually easier to measure a line than a point, and that is such a common approach that these maxima are called *spectral lines*. The telescope part of the spectroscope allows you to scan through angles to find these lines and precisely measure their θ angle. It has an adjustable eyepiece, to allow you to focus on the lines, and cross-hairs for visual alignment. The angular position of the telescope can be measured with scales situated on either side of the spectrometer. They measure in degrees with Vernier scale markings down to one arc minute.

Use of a Vernier scale takes a bit of explanation and some practice. You can easily find an extended description online, here we'll introduce it with brief examples. Two sets of scales are arranged so that they slide past each other. The outer one has lines marking for each half degree, while the inner scale has markings corresponding to 0-30 arc minutes. The two scales are used together. Cartoons of two different measurements are shown in Fig. 20 to illustrate the procedure. First you look where the zero mark on the inner scale is located. In the left hand cartoon that is just past $210°$; on the right side it



210° 15'

229° 46'

Figure 20: Examples of reading a Vernier scale. The middle photo is from the spectroscope that you will use. The left and right cartoons are discussed in the text to illustrate how to read the scale; the arrows in the cartoon show where the lines on the inner and outer scale are most closely aligned for these examples.

is just past $229.5°$. Then you look where the lines on the inner and outer scales align most closely. In the left hand cartoon that happens for the $15'$ mark; on the right side it is the $16'$ mark. These measure the number of arc minutes to add to the values obtained from the first step. So, the measurement on the left side is $210° + 15' = 210° + (15/60)° = 210.25°$, and the measurement on the right is $229.5° + 16' = 229.5° + (16/60)° = 229.77°$. To practice this yourself, figure out the measurement for the photo in the middle and compare it to what your lab partner gets.

The uncertainty on θ is driven by the one arc minute granularity of the scales, so it is appropriate to use $\delta\theta = 1'/\sqrt{12} = 0.3' = 0.005°$. That provides rather precise angular measures, but it is possible to make mistakes in reading or calculating the values. There are a few things you can do to reduce the chance of such mistakes and their impact. First of all, both you and your lab partner should record the values and compare. Secondly, the spectroscope has a scale on both sides, where the angles differ by $180°$; you should record both. If, after correcting for the $180°$ difference you find that their values differ by more than the $0.005°$ precision of the scales, then there is an additional source of uncertainty. That could arise from things like warping of the scales, and it would then be the dominant uncertainty on your θ measurements. This approach of making multiple, independent measurements of the same quantity is always a good idea to help you validate your results.

You should also record the raw readings in degrees and arc minutes, and then calculate the corresponding decimal degree values. This will help you disentangle any calculation errors that may occur. While the calculations in this case are not complicated, and the consequences of a mistake are small, this sort of fastidious data recording is a good habit to develop for your later research work, when both the chances and consequences of mistakes increase.

Finally, beware that the angle scale just measures an angular position on some arbitrary scale. What you want to know for the wavelength determination in $d \sin\theta = \lambda$ is the angle *between* maxima. So, you need to measure two angles and subtract, just like measuring a length requires two position measurements. You could do that by measuring the angular position of the central maximum, which we can call θ_0, and the angular position of the $m = +1$ maximum, which we can call θ_{+1}. Then $\lambda = d \sin\theta_{+1} - \theta_0$. Similarly, you could measure the angular position of the $m = -1$ maximum and use $\lambda = d \sin\theta_0 - \theta_{-1}$. If everything is perfectly aligned, $\theta_{+1} - \theta_0 = \theta_0 - \theta_{-1}$. If not, then neither value will be correct. You can cancel the effects of any mis-alignment by using *both* the $m = +1$ and $m = -1$ positions, i.e., $\lambda = d \sin(\theta_{+1} - \theta_{-1})/2$. Since you can't be certain that the diffraction grating is exactly perpendicular to the light path and the spectroscope's angular coordinate system, you *must use this difference of the two angles*. Finally, note that the direction in which the angles increase or decrease on the scale is arbitrary, so you should use the absolute value of the differences, i.e.,

$$\lambda = d \sin|\frac{\theta_{+1} - \theta_{-1}}{2}| \tag{10}$$

5.1.2 Safety and care of equipment

Be careful not to touch the glass part of the lamps. It can get hot, and the oils on your fingers can cause the glass to overheat and shorten the life of the lamp.

To help you discern the dim spectral lines, it is best to leave the room lights off. There is a desk lamp, but use it only sparingly. You can use the small flashlight provided when you need to record the angle scales. The darkness and crowded room means you might back into or trip others who are trying to move around the lab (like your TA). In such situations it is a good idea to just move a bit more slowly.

The spectroscope has clamps that will hold it in place while you measure angles. If you forget to unclamp things before moving, it can cause damage. So, do not force any motion that doesn't feel smooth.

5.2 Getting started and gaining familiarity

You want to start with the hydrogen discharge tube. There are two other tubes that contain other elements.

Play around with the spectroscope to make sure you understand how it works. You should align the telescope and collimator and adjust the slit and crosshairs.

The grating should be approximately perpendicular to the light path. You can adjust it and lock the table in place. Any misalignment will skew the angles slightly, so as noted above, you should measure the position of the spectral lines for both $m = +1$ and $m = -1$ and use half the difference to reduce the effect of any misalignment.

Look for each of the emission lines. You should see a total of four, although the last one, which is violet, is dim and difficult for some people to see. You can increase the slit width to make the lines easier to see, although you want to balance the need for visibility with the desire for narrow lines that can be measured more precisely. You may also see a background of other colors besides the main bright lines. This background can come either from stray light hitting the diffraction grating or from the ultraviolet light emitted by the hydrogen lamp exciting other atoms to emit. In fact, that is how the commonly used flourescent lights work. Ultraviolet emission produced by running a current through the gas in the long tubes excites atoms in the white flourescent material that coats the inside of the tube. That material emits a with a large number of different wavelengths that we perceive as white light. You can use the hand held spectroscope (black trapezoidal object) provided to check out the spectrum of the flourescent lights in the lab's ceiling.

5.3 Measuring the spectrum and extracting Rydberg's constant

Measure the angles for each of the spectral lines. (Remember to measure the position on both sides to correct for any misalignment of the grating, and to clamp the telescope before making a reading and then unclamp it before moving again.) It is appropriate for you and your lab partner to take turns looking through the telescope. In fact, I suggest that one person collect a full set of measurements and the other collect another full set. You can each record readings for the other's dataset as a cross-check, but it is instructive to independently collect and analyze your own dataset.

Once you have measured the angular positions of each of the emission lines, calculate their wavelengths and determine the uncertainty on the wavelengths. To find the uncertainty you will need to propagate the uncertaintes on each of the measured angles, $\delta\theta_i$, to an uncertainty on λ using the standard error propagation approach described in Equation (1).

Then use Equation (9) to extract the Rybderg constant, R. You can do that in several ways, such as averaging all the measurements or fitting them. A fit is usually the best approach because it incorporates the uncertainties on each of the measurements, and the χ^2 of the fit provides a quantitative way to validate the measurements and uncertainties. So, fit your data.

Think about what functional form is the best to use for the fit. It is not simple to solve Equation (9) for $\lambda(n)$, but you can calculate things that are easy to plot, e.g., plotting $y = 1/\lambda$ vs $x = 1/2^2 - 1/n^2$ should be a line with a slope equal to R. Remember to propagate the uncertainties properly though, the uncertainty on y, i.e., on $1/\lambda$ can be found from the uncertainty on λ using Equation (1).

If the χ^2 of your fit is too large or too small, it can indicate that your uncertainties are incorrectly estimated or that you just made a mistake in calculating. The best way to figure this out is to make multiple measurements and compare them. If different measurements of the same quantity disagree completely then there is a mistake somewhere. (Note that recording the "color" of the lines you have observed provides a coarse measurement of their wavelength; use that to validate that you have not made a major mistake in calculating λ from the measured angles). You can also compare your measured values from one set of data to the values from another set. They should agree within approximately the uncertainties that you have determined. If the disagreements are a bit larger than the uncertainties, then you may have simply underestimated the uncertainties. Increasing the uncertainties on the λ's to make the level of agreement between repeated measurements of them match a 68% confidence level will then make the uncertainty on the fitted R value more appropriately correspond to a 68% confidence level.

5.4 Measuring the spectrum of an unknown source

There are two other discharge tubes provided. They contain other elements. Replace the hydrogen tube with either one of these and measure its spectrum. You can remove the tubes by pushing them up and out of the black mounting box. The new ones should then slide in easily. Don't force it.

Compare your measured spectrum to those given for standard sources in the following pages and determine the composition of the discharge tube. Note that in some cases, the relative intensity of the different lines can be helpful, so it is wise to record a qualitative measure of that.

5.5 Going Beyond

There are several ways that exceptional students could go beyond the basics in this lab. I list a few ideas below, but you are encouraged to use your own creativity.

- Can you think of a way to quantitatively record the relative intensity of the emission lines?
- Can you use your measured value of the Rydberg constant, and Bohr's energy level calculation, to extract a limit on the proton mass?
- Now that you understand how to use the spectroscope, can you use it to measure the spectrum of something else.

Pen-Ray® Rare Gas Lamp Spectra

UVP

ULTRA-VIOLET PRODUCTS

CONTENTS

PEN-RAY® Rare Gas Lamp Spectra

It is a pleasure to present in the following pages a report that fills a much needed gap in the present knowledge of spectra of the rare gases. We are indebted to Dr. Wilbur Kaye, Research Director of Beckman Instruments, Inc., Fullerton, California, for his interest in the rare gas forms of the PEN-RAY lamps. We thank Dr. Kaye and Beckman Instruments for their generosity in allowing us to publish the report and spectra.

Preamble and comments on Dr. Kaye's report

PEN-RAY lamps have been manufactured for over 25 years in the standard Argon-mercury quartz lamp type. Because of remarkable stability, low operating temperature, and good, discrete spectra; PEN-RAY lamps have received wide recognition. Thousands are in use for calibrating various photometric instruments or as light sources in optical instrumentation.

Increased emphasis on aerospace studies have developed interest in modifications of the PEN-RAY lamps, relative to special gas fillings, sizes, and shapes of lamps. As many of the requirements were for noble gas lamps, the various rare gases have been made available as standard equipment.

The initial rare gas lamps, made available to several scientists for testing, caused considerable interest. The lamps were then displayed at three large gatherings of scientists with very favorable reactions.

The most important consideration pointed out by the various groups was the need for spectra of the lamps. The report and spectral data presented here is a result of this expressed need.

The PEN-RAY lamps were investigated with a Beckman DK-U (Universal) Prism-Grating Spectrophotometer. This versatile instrument was described by Dr. Kaye in *Applied Optics,* Vol 2, p. 1295, Dec. 1963, under the title, "A Universal Spectrophotometer."

In the report on the PEN-RAY® lamps which follows, the 10 ma power supply is mentioned as a probable compromise. It is true that lamps vary slightly in manufacture and that pure gas lamps are particularly sensitive in their current requirements. The investigator may find it desirable to operate the power supply from a variable transformer in order to make adjustments if a lamp does flicker slightly.

The Neon-mercury lamp, which was investigated, is not a rare gas lamp as under normal conditions the mercury spectrum is superior to the Neon spectrum. When the lamp is cooled sufficiently, the Neon spectrum becomes superior. However, it is interesting to note that several mercury spectral lines are stronger from the Neon-mercury lamp than from the standard Argon-mercury lamp.

As noted in the report the spectral lines of Xenon are relatively weak in a low pressure discharge lamp. These rare gas lamps, though they do not produce a continuum like high pressure arc lamps, are more valuable for certain applications in that they do produce discrete, sharp spectral lines.

The two basic areas of application are:

I. Reference Standards
II. Basic Research

As reference standards the lamps may be used in several areas.

1. Wavelength calibration and alignment of optical instruments, i.e., Spectrophotometers.
2. Calibration of spectrograph spectra by imposing spectra from the lamps onto the film.
3. Evaluation of performance, i.e., resolution of spectrophotometric instruments.
4. Checking output intensity of optical instrumentation.
5. Spectrochemical analysis with colorimeters.
6. Checking sensitivity of detectors as photocells, photomultipliers and thermopiles at specific known wavelengths.

In basic reseach some of the fields of interest may be:

1. Basic study of the emitted light itself.
2. Study of the light after passing through a system or media.
3. Reflection and irradiation studies with various monochromatic spectral lines.
4. Use of the light sources as devices for interferometric measuring of thin films and surface finishes. Also, for studying characteristics and structure of the light itself.

The PEN-RAY lamps used for the accompanying spectra were the basic 11SC-1 shape with a lighted length of approximately two inches. Other models with shorter lighted lengths are possible, as well as models of smaller and larger diameter. Single bore lamps can also be made in various configurations as required.

Emission Spectra of Argon, Helium, Krypton, Neon and Xenon PEN-RAY® Lamps

by Wilbur Kaye, Beckman Instruments, Inc.

Introduction

A number of new PEN-RAY lamps containing various gases was placed at our disposal by ULTRA-VIOLET PRODUCTS INC., San Gabriel, California. These lamps are of a convenient size and power for general laboratory use. The PEN-RAY lamp containing mercury vapor has long been used for alignment and wavelength checkout on spectrophotometers. The spectra of these lamps, obtained with a Beckman DK-U prism-grating spectrophotometer, are reproduced here.

Lamps

PEN-RAY lamps were filled with the following gases: argon, helium, krypton, neon, xenon and a mercury-neon mixture. The lamps have a dual bore fused silica tube 2⅛ (5.39cm) inches long confining the arc. The lamps are powered by a small transfomer-ballast (SCT-3) providing the proper starting potential at 270 to 1000 volts at 10.0 ma during lamp operation. Total power consumption is 14 watts. The ideal power requirements of these lamps no doubt varies with the filling gas and the use of one power supply is a compromise. In particular, the neon and helium lamps flickered excessively and it was necessary to raise the primary voltage to 130 volts with a Variac to obtain acceptable stability.

The lamp containing the mercury-neon mixture exhibited an emission spectrum that was very sensitive to lamp temperature. At low temperatures (immediately after starting or when cooled with a jet of air) the neon spectrum is favored. At normal operating temperature the mercury emission predominates. Undoubtedly, this effect is due to the vapor pressure of mercury in the lamp.

Model 11-SC-1
PEN-RAY Lamp

Spectra

Spectra were obtained on these lamps using the DK-U spectrophotometer and are found in figures 1-0. Both prism and grating monochromators were used for scanning the lamp emission above 3200 A and the prism alone was used for shorter wavelength scans. The RCA tri-alkali detector C70101 was used for all spectra employing the grating monochromator and the RCA 7200 detector was used when the prism monochromator alone was employed. The amplifier gain and dynode voltage was varied to give the best spectra. A pen response time of 0.1 second was used for all spectra. The lamps were mounted in a high pressure xenon backplate (W128240). The slit width labeled on the figures refers to the grating monochromator except where "prism only" is noted. When both monochromators were employed, the slit width of the prism monochromator was kept at 2.0 mm.

In regions of intense emission, the spectra were run at two different slit widths in order to display both strong and weak likes. In general the wavelength scale has been expanded for the spectra obtained at the narrower slit widths. Wavelength intervals between 2000-8000 A containing no lines of appreciable intensity are not reproduced.

Wavelengths of the lines are given to the nearest 0.1 A. These values are taken from the 41st edition of the Handbook of Chemistry and Physics. Doubtful lines have been identified by comparison with the hydrogen emission spectrum. All wavelengths are given in standard air, except the mercury lines below 2000 A and these are corrected for nitrogen at 25° and 760 mm.

The intensities of the emission lines bear little relationship to the Handbook values. Apparently the excitation conditions are different. Close examination reveals that the emission from the PEN-RAY lamps arises almost exclusively from neutral atoms. In some cases numerous expected lines were completely absent. For example, the ultraviolet lines of argon are almost completely absent. The emission intensity from the xenon lamp appears unusually low and may be due to a faulty lamp.

The intense emission lines from these lamps are listed by wavelength in table I. These lines should be most useful.

The spectrum of the mercury-neon mixture is given only under the condition of a cooled lamp. The uncooled lamp exhibits very weak neon lines, hence, is essentially identical with a pure mercury lamp.

Useful Wavelengths With PEN-RAY® Lamps

Table I

(Air or N_2)	Lamp	(Air or N_2)	Lamp	(Air or N_2)	Lamp
		4319.6	Kr		
1849.1	Hg	4333.6	A	6717.0	Ne
1941.7	Hg	4358.4	Hg	6929.5	Ne
2262.2	Hg	4362.6	Kr	6965.4	A
2378.3	Hg	4376.1	Kr	7024.1	Ne
2482.0	Hg	4387.6	He	7032.4	Ne
2536.5	Hg	4453.9	K	7059.1	Ne
2652.0	Hg	4463.7	K	7065.2	He
2803.5	Hg	4471.5	He	7067.2	A
2893.6	Hg	4502.4	K	7173.9	Ne
2967.3	Hg	4713.1	He	7245.2	Ne
3021.5	Hg	4921.9	He	7272.9	A
3125.7	Hg	5015.7	He	7281.4	He
3131.7	Hg	5047.7	He	7384.0	A
3187.7	He	5330.8	Ne	7438.9	Ne
3341.5	Hg	5341.1	Ne	7488.9	Ne
3369.9	Ne	5400.6	Ne	7503.9	A
3417.9	Ne	5460.7	Hg	7514.6	A
3447.7	Ne	5562.2	K	7535.8	Ne
3466.6	Ne	5570.3	K	7544.1	Ne
3472.6	Ne	5769.6	Hg	7587.4	K
3520.5	Ne	5790.7	Hg	7601.5	K
3593.5	Ne	5852.5	Ne	7635.1	K
3650.2	Hg	5870.9	K	7685.2	K
3654.4	Hg	5875.6	He	7694.5	K
3663.3	Hg	5881.9	Ne	7723.8	A
3888.7	He	5944.8	Ne	7854.8	K
3949.0	A	5975.5	Ne	7948.2	A
3964.7	He	6030.0	Ne	8006.2	A
4026.2	He	6074.3	Ne	8014.8	A
4044.4	A	6096.2	Ne	8059.5	K
4046.6	Hg	6143.1	Ne	8103.7	A
4077.8	Hg	6163.6	Ne	8104.4	K
4120.8	He	6217.3	Ne	8112.9	K
4158.6	A	6266.5	Ne	8115.3	A
4164.2	A	6304.8	Ne	8190.1	K
4181.9	A	6334.4	Ne	8231.6	Xe
4191.0	A	6383.0	Ne	8263.2	K
4198.3	A	6402.3	Ne	8264.5	A
4200.7	A	6506.5	Ne	8280.1	Xe
4259.4	A	6532.9	Ne	8298.1	K
4272.2	A	6599.0	Ne	8377.6	Ne
4274.0	Kr	6678.2	He	8408.2	A
4300.1	A	6678.3	Ne	8424.6	A

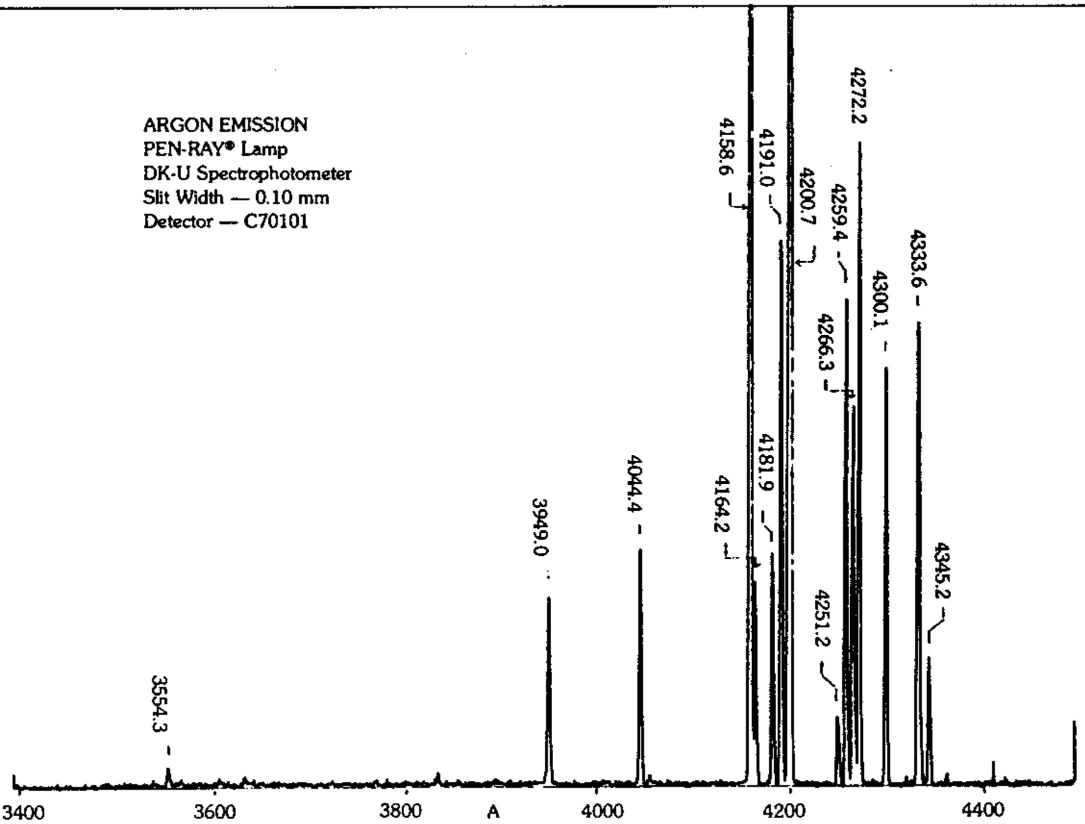


FIGURE 1

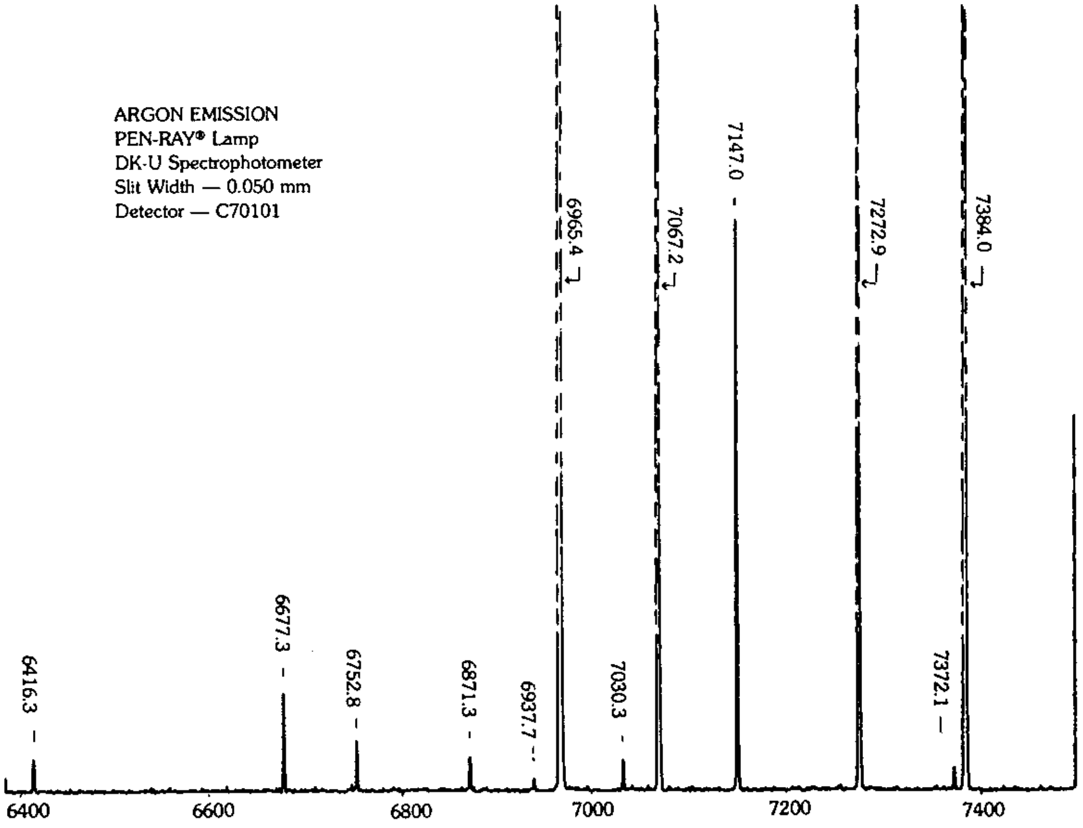


FIGURE 2

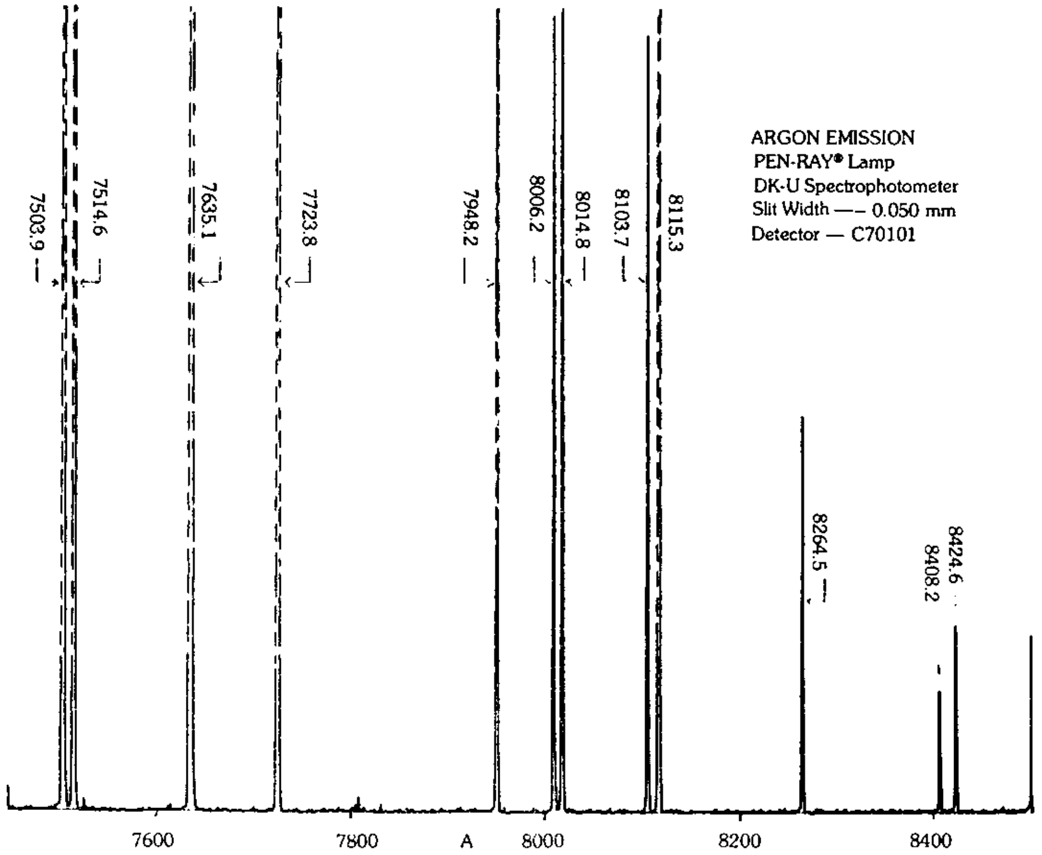


FIGURE 3

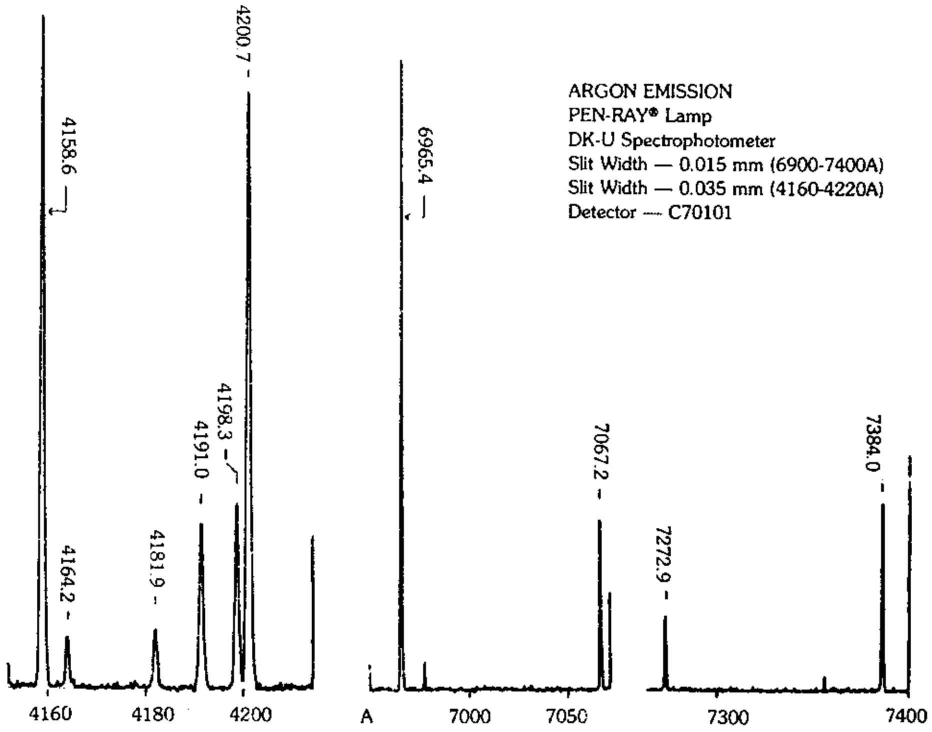


FIGURE 4

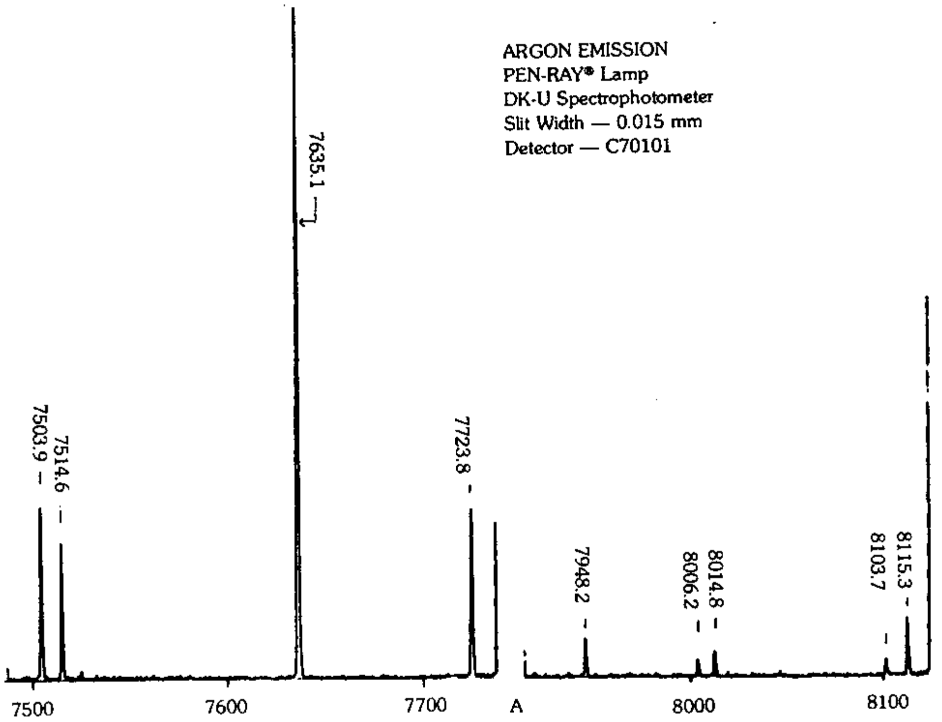


FIGURE 5

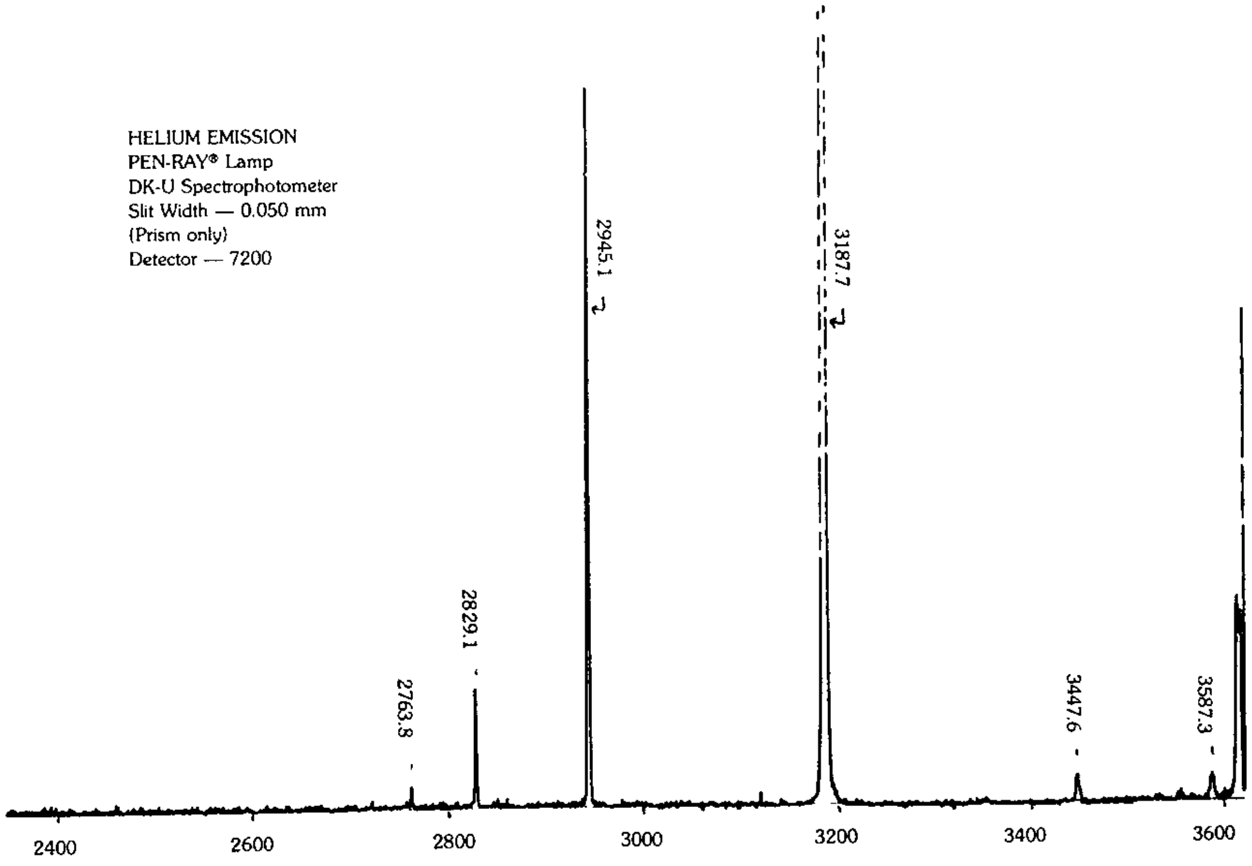


FIGURE 6

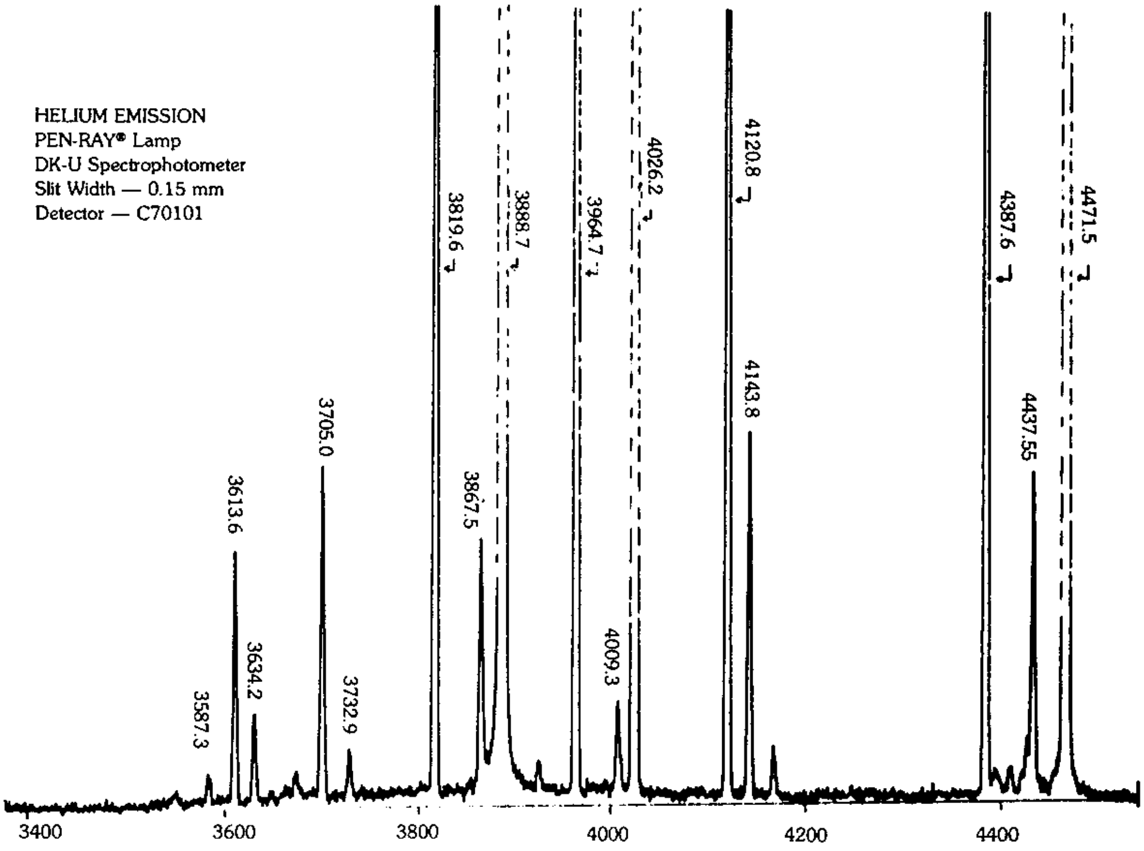


FIGURE 7

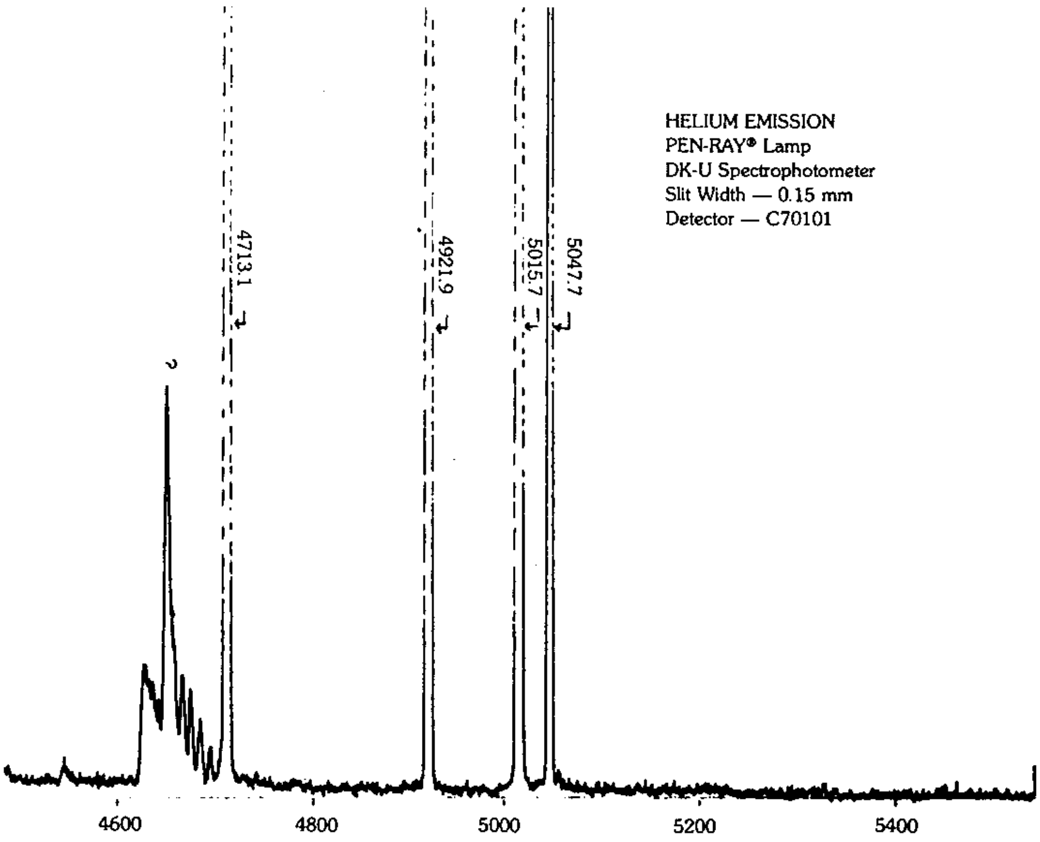


FIGURE 8

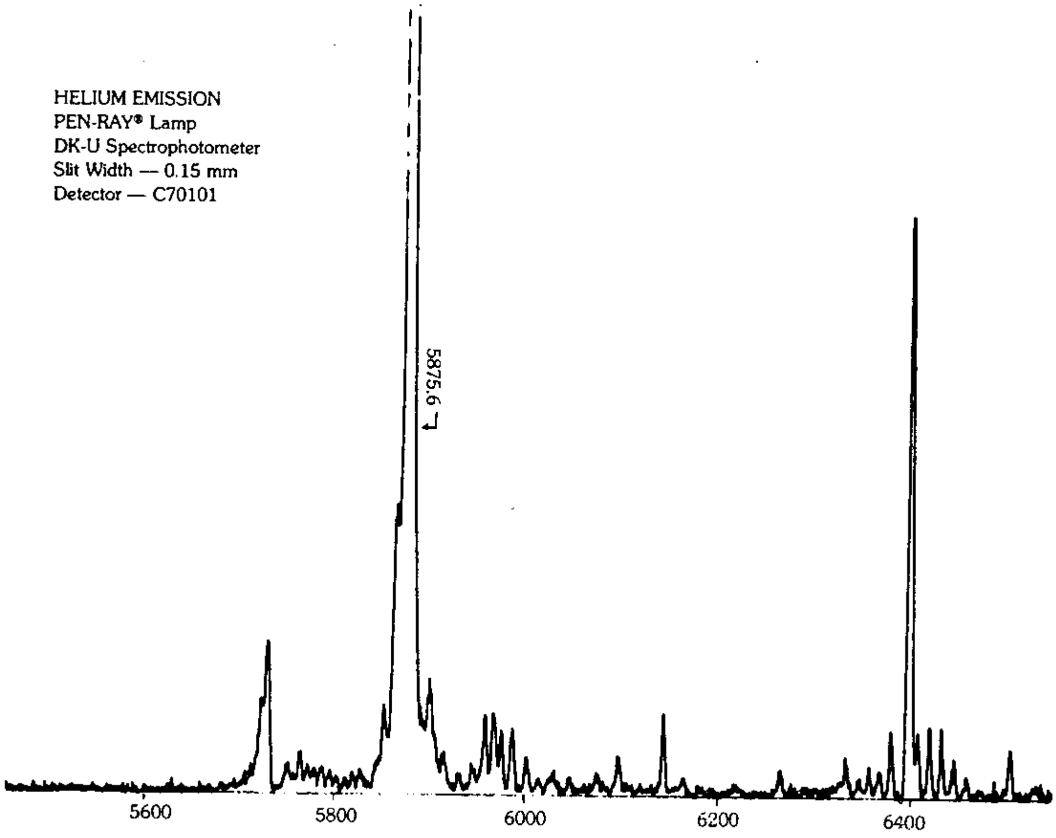


FIGURE 9

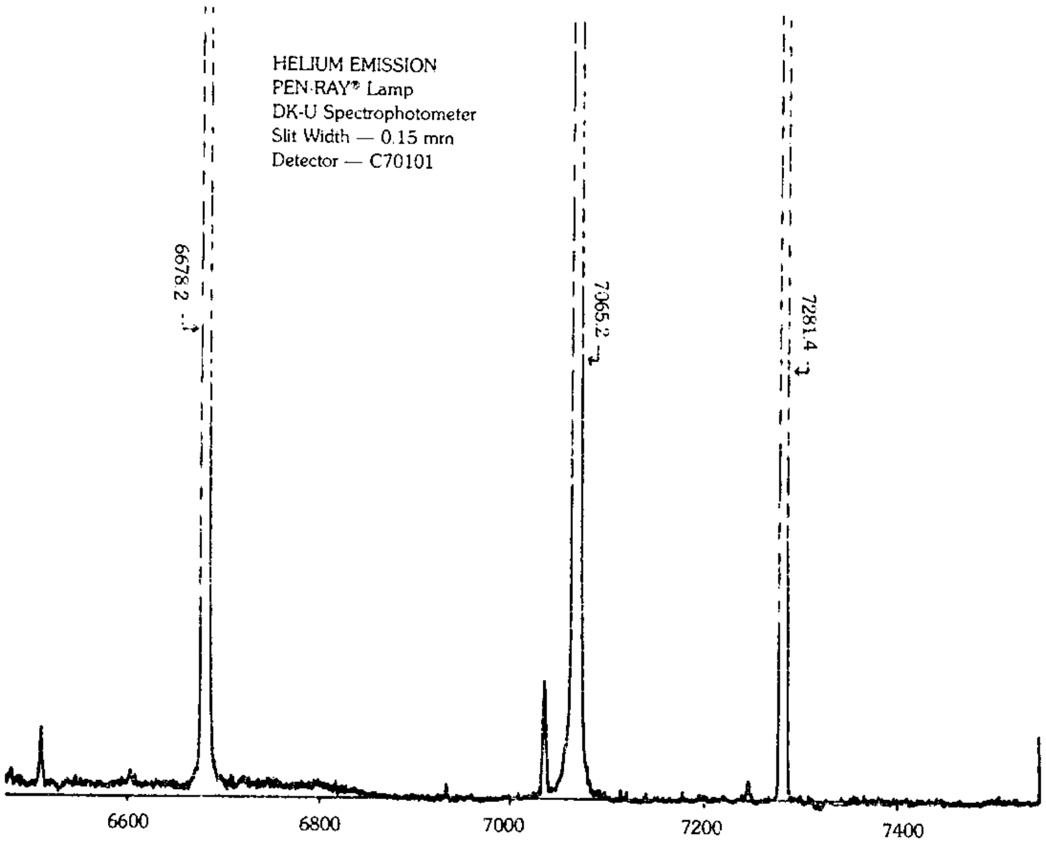


FIGURE 10

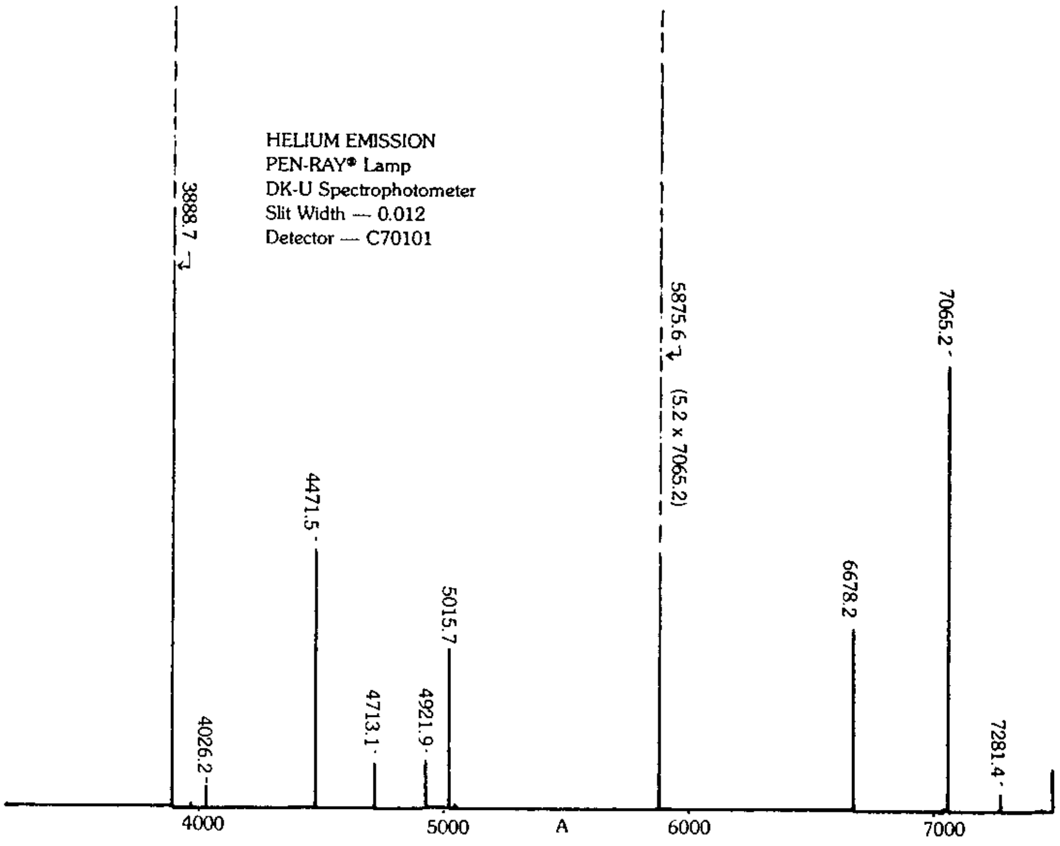


FIGURE 11

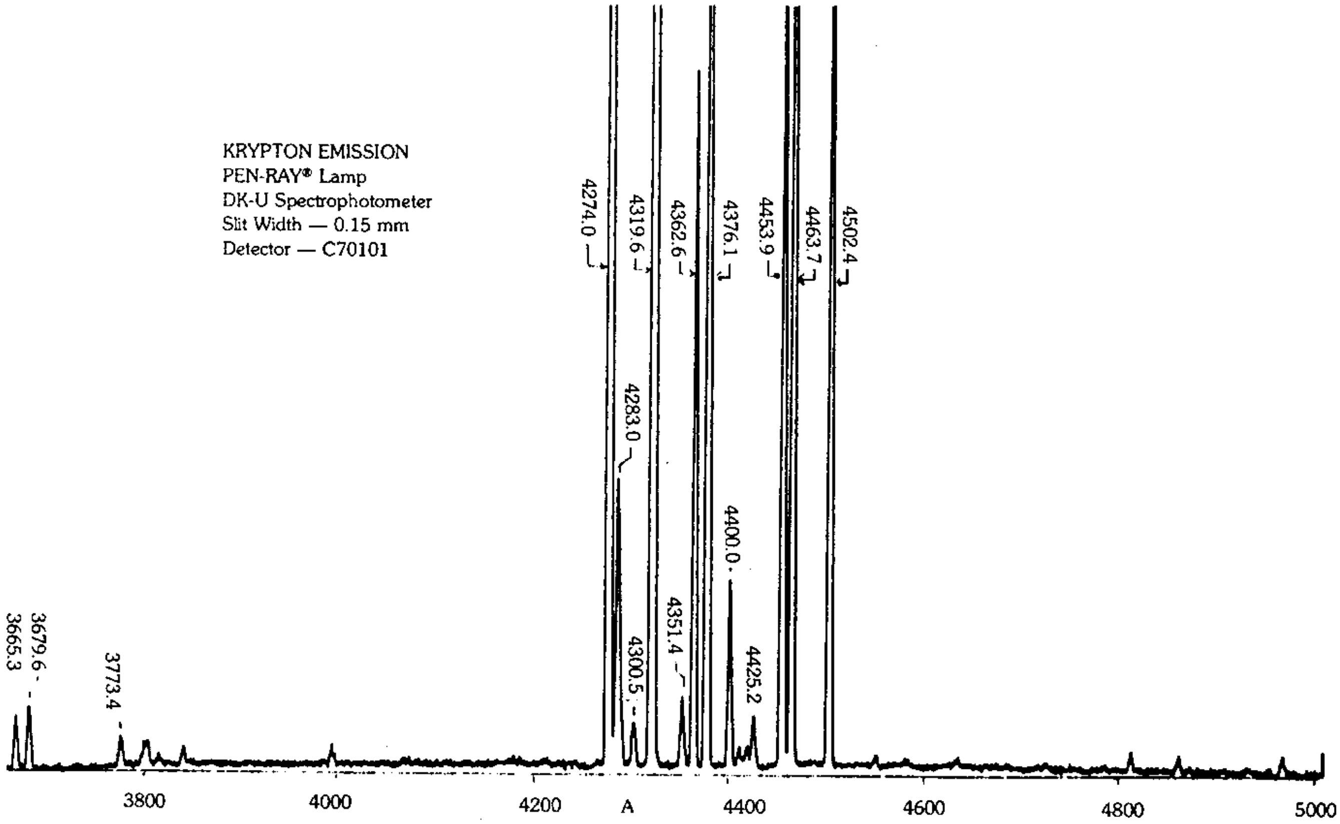


FIGURE 12

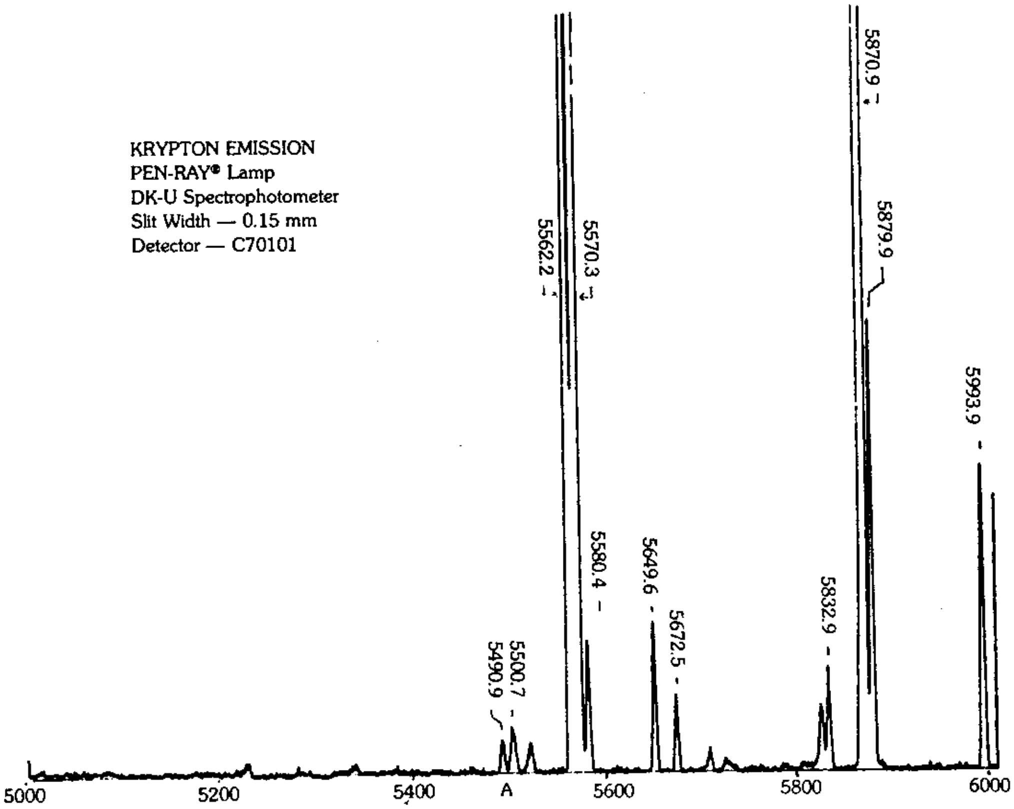


FIGURE 13

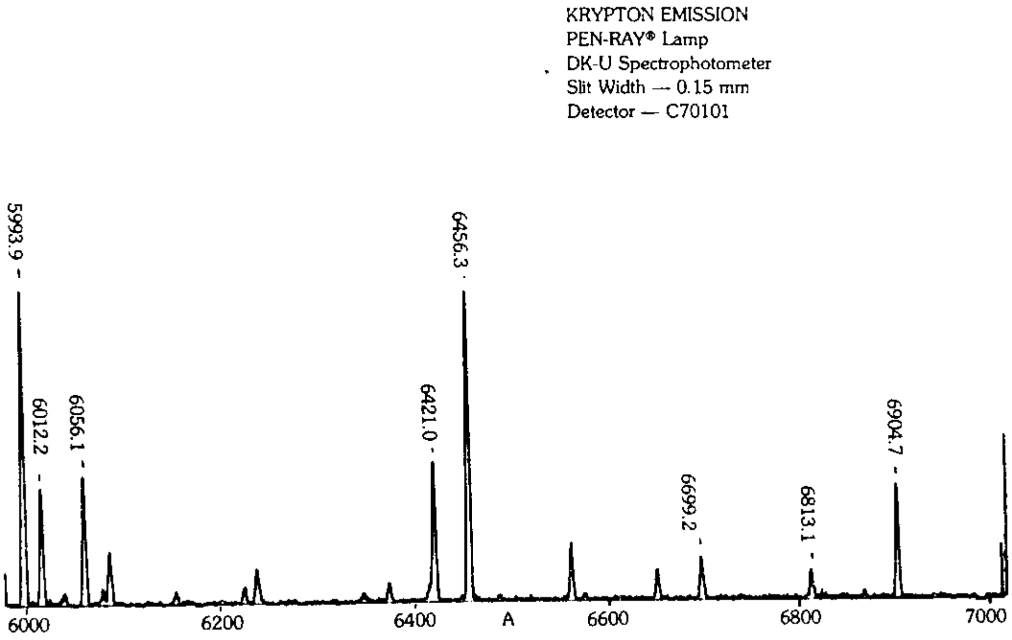


FIGURE 14

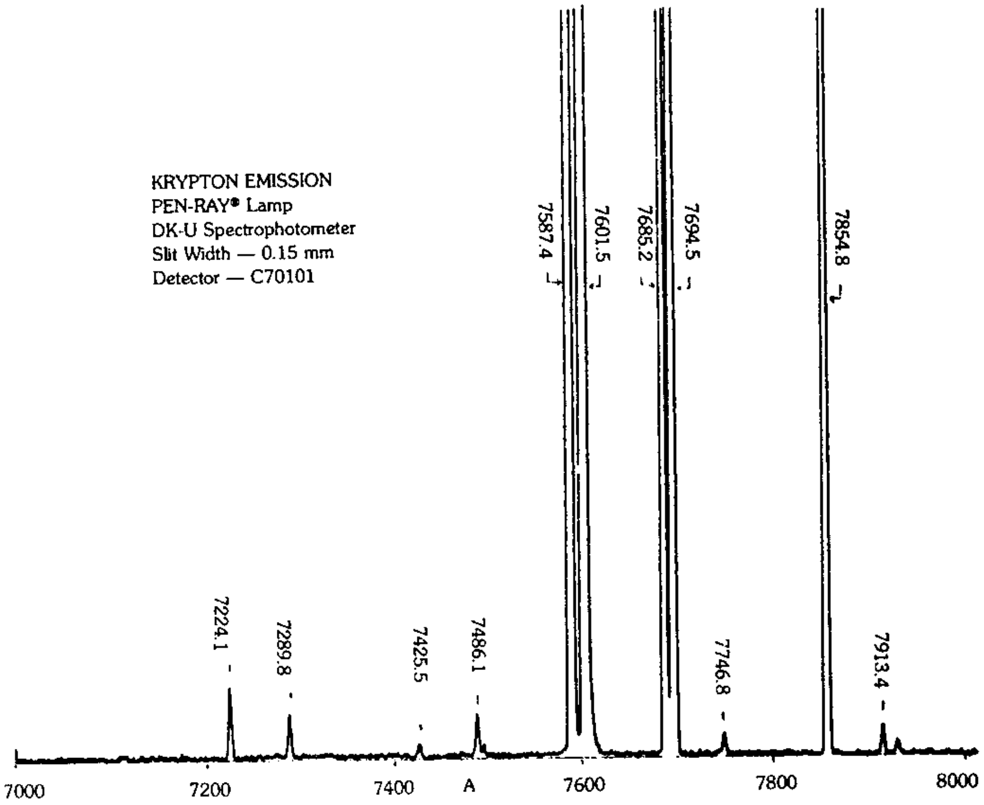


FIGURE 15

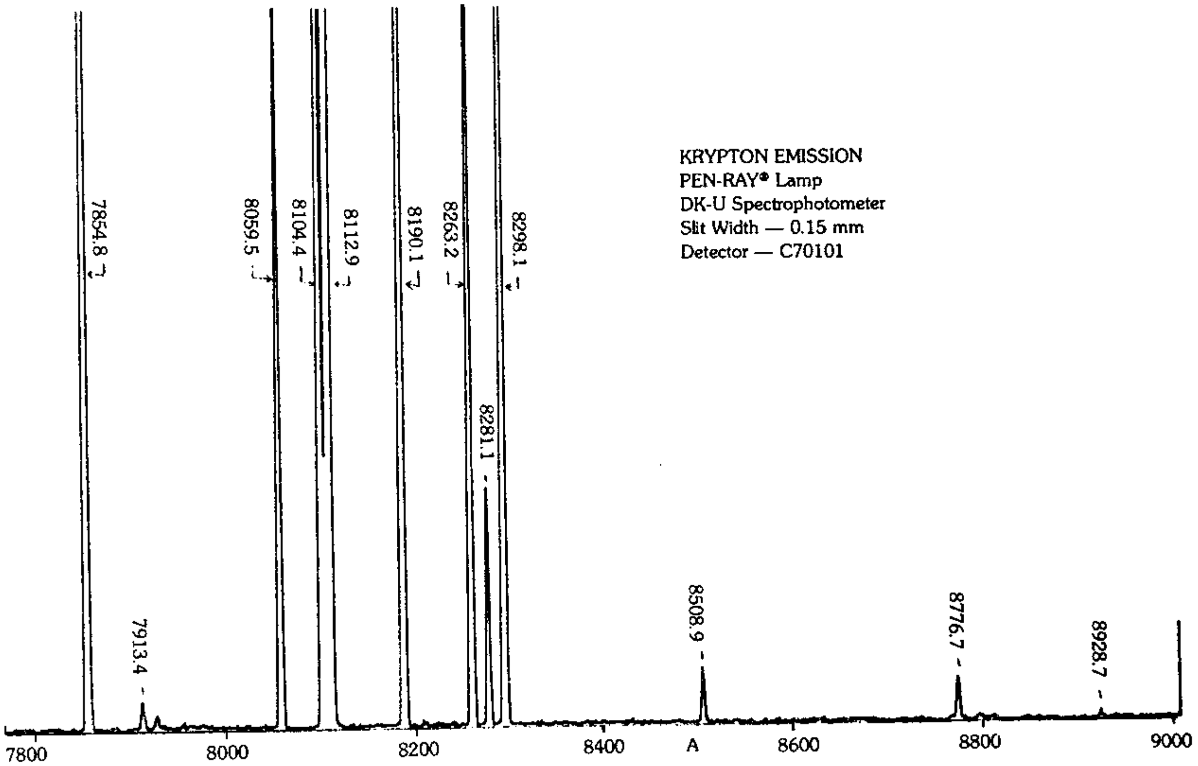


FIGURE 16

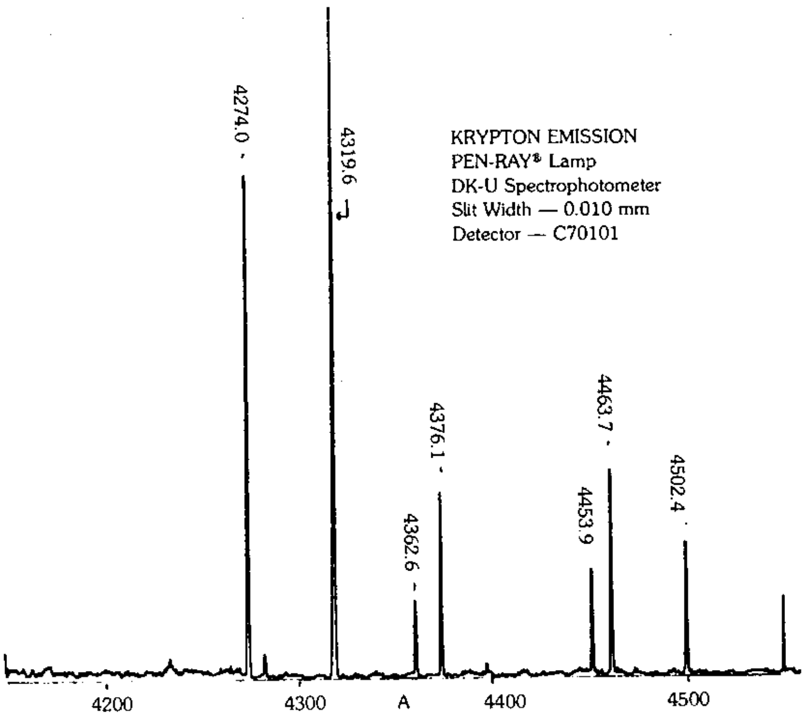


FIGURE 17

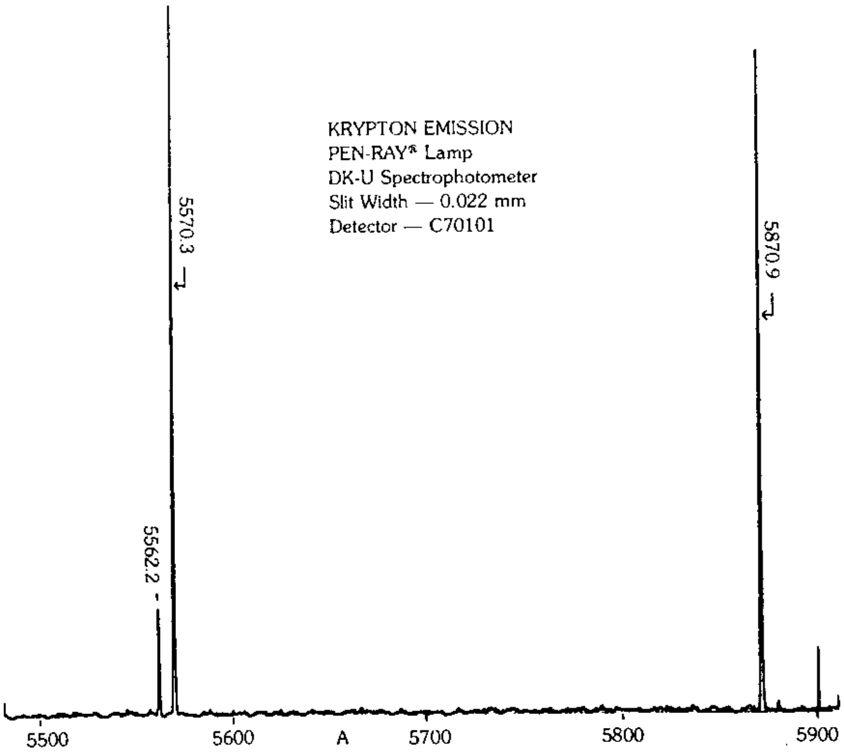


FIGURE 18

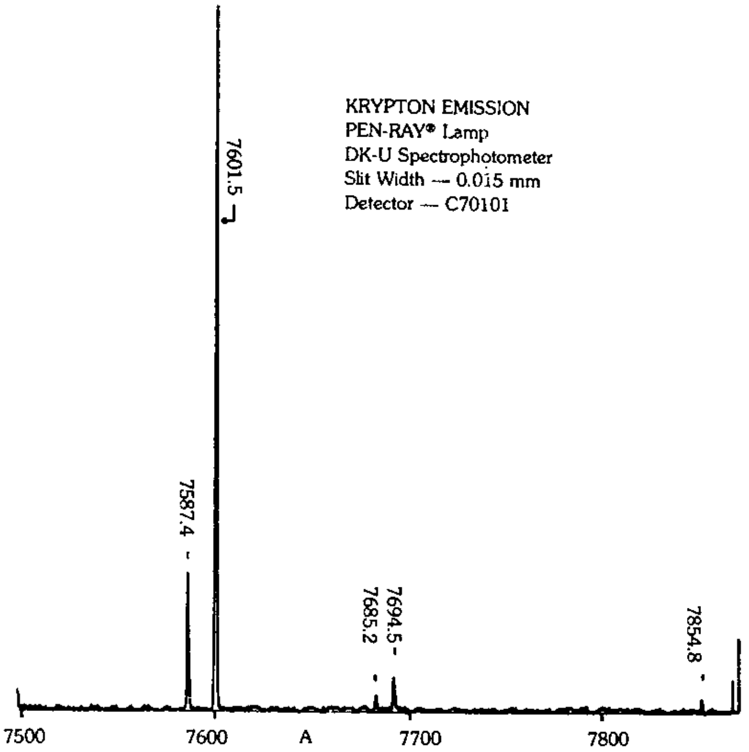


FIGURE 19

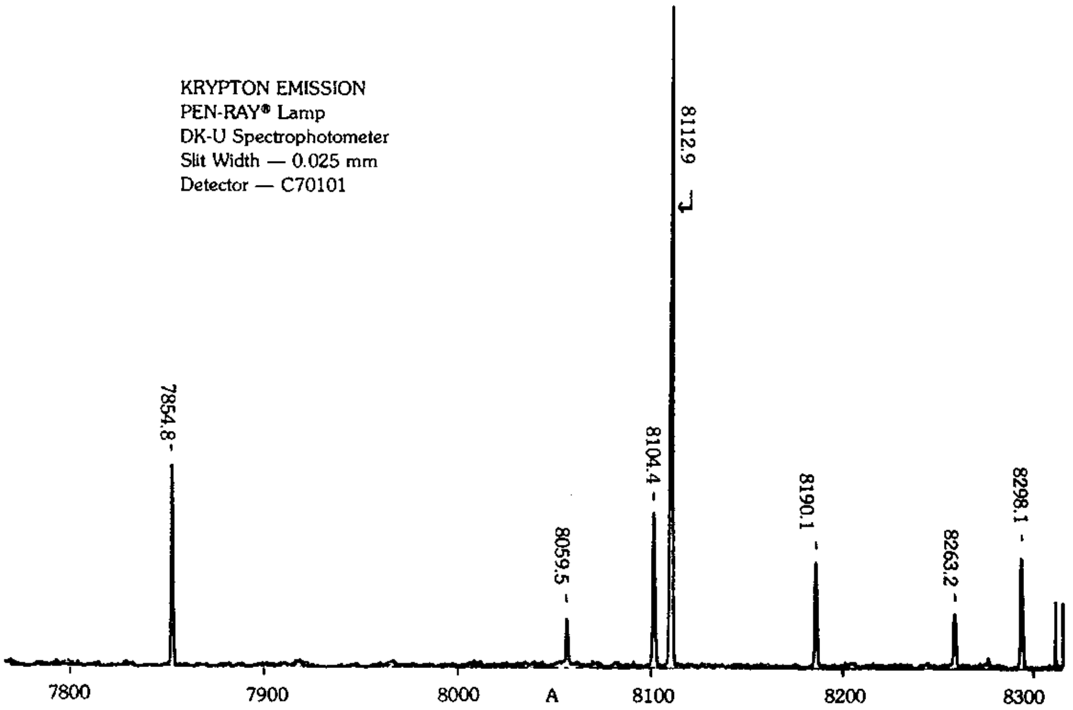


FIGURE 20

NEON EMISSIN
PEN-RAY® Lamp
DK-U Spectrophotometer
Slit Width 0.10 mm (2nd order)
Detector 1P28

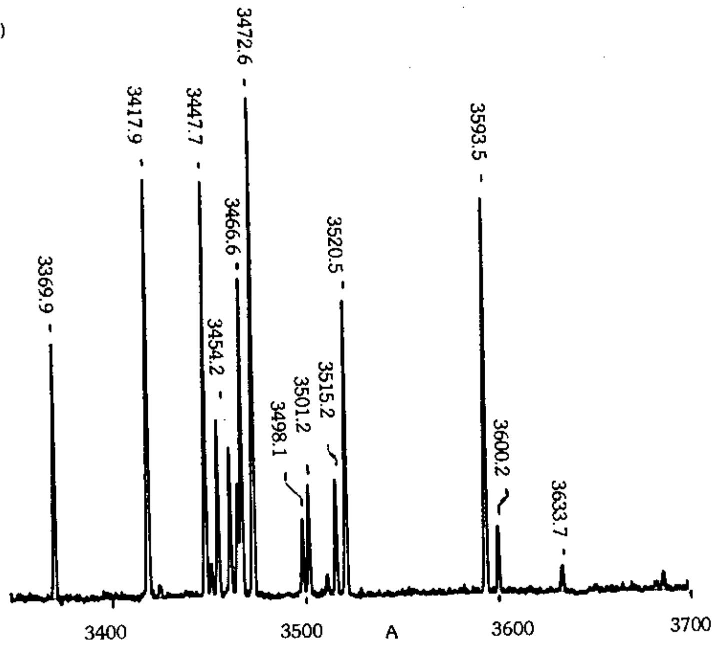


FIGURE 21

NEON EMISSION
PEN-RAY® Lamp
DK-U Spectrophotometer
Slit Width — 0.050 mm
Detector — C70101

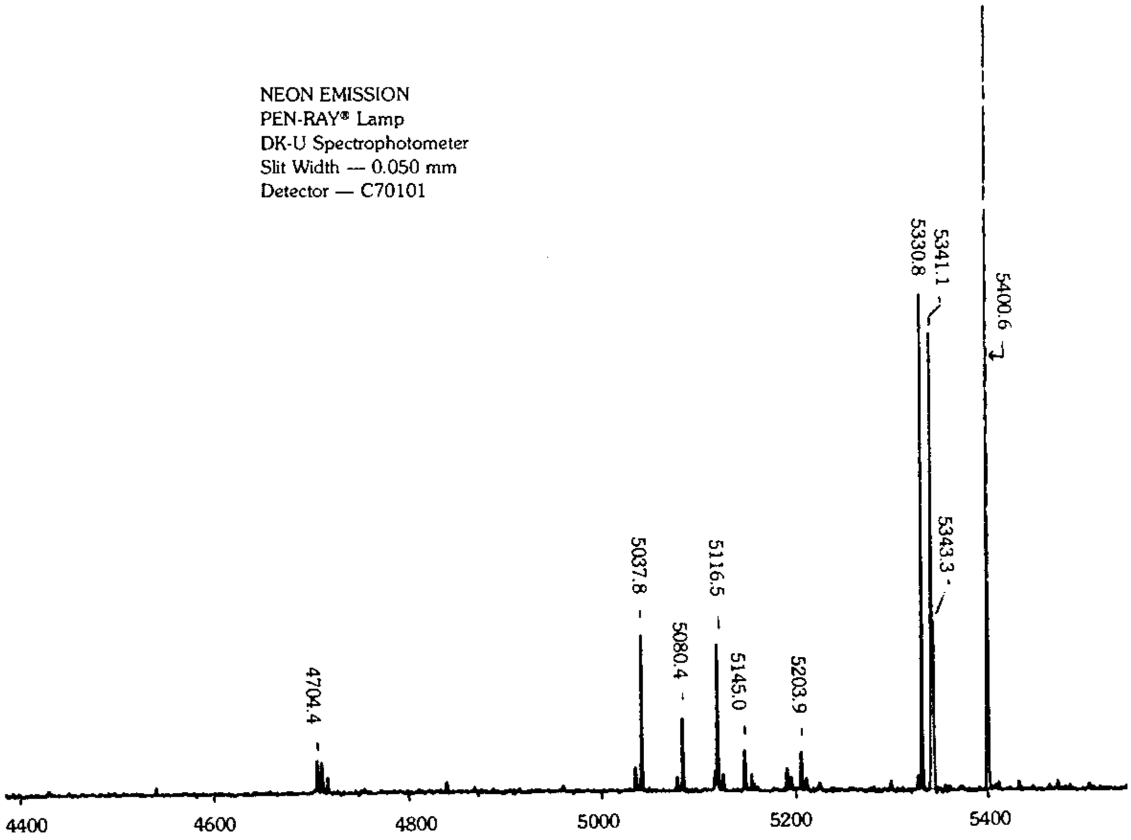
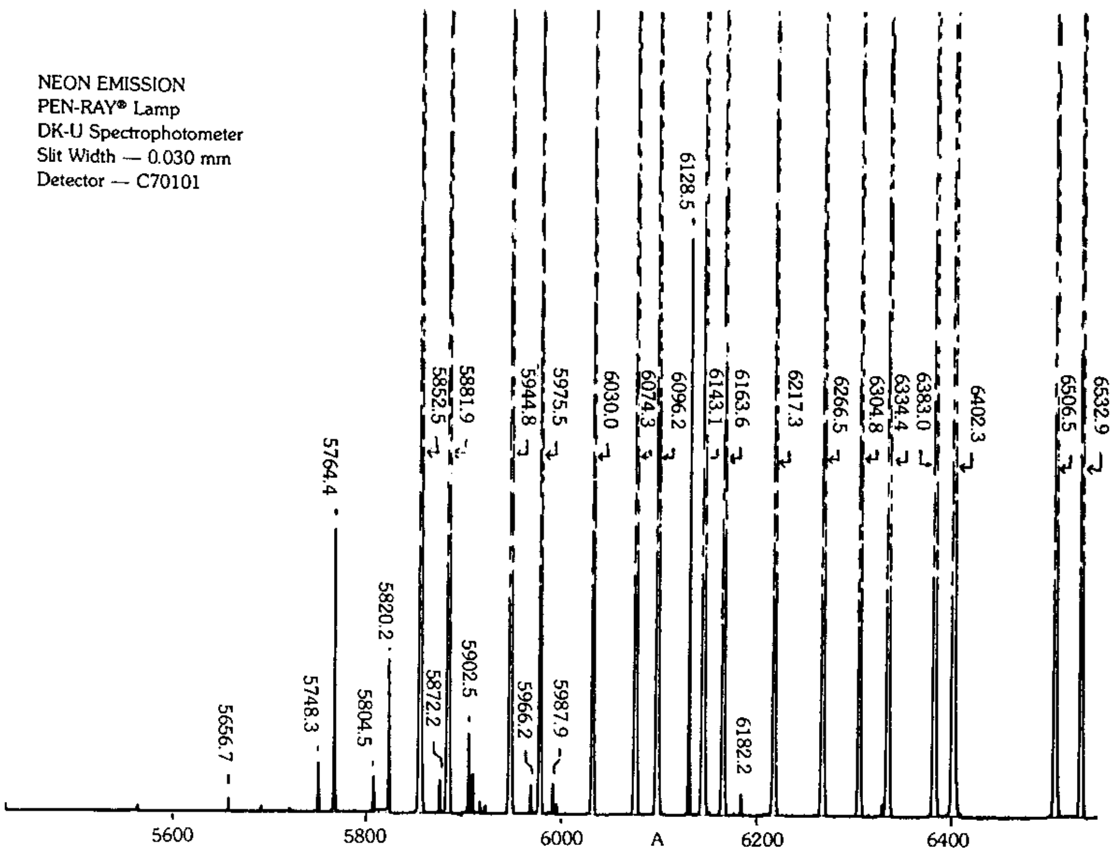
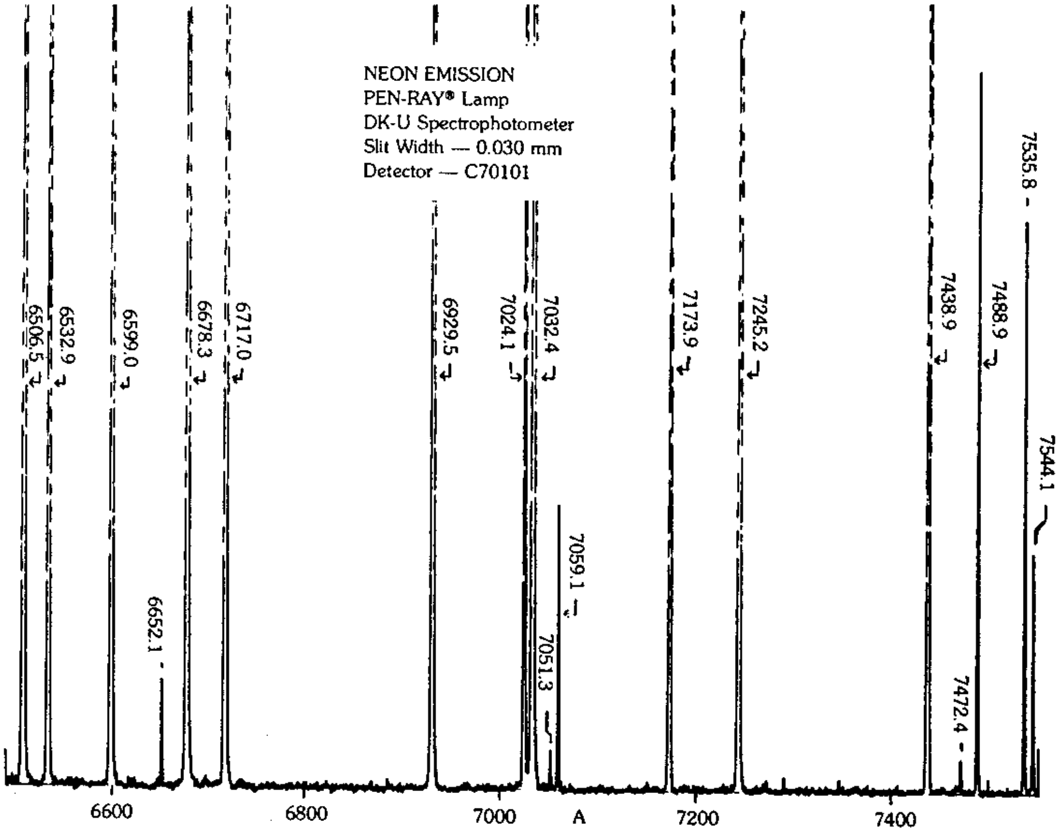


FIGURE 22



NEON EMISSION
PEN-RAY® Lamp
DK-U Spectrophotometer
Slit Width — 0.030 mm
Detector — C70101

FIGURE 23



NEON EMISSION
PEN-RAY® Lamp
DK-U Spectrophotometer
Slit Width — 0.030 mm
Detector — C70101

FIGURE 24

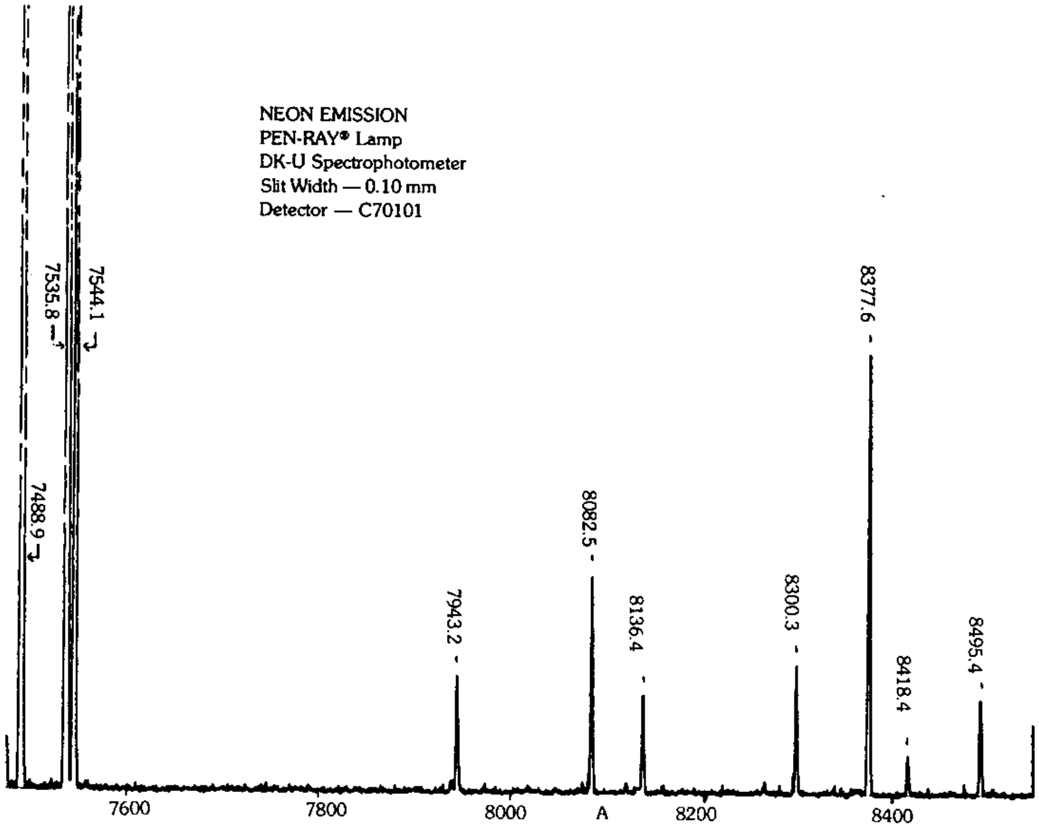


FIGURE 25

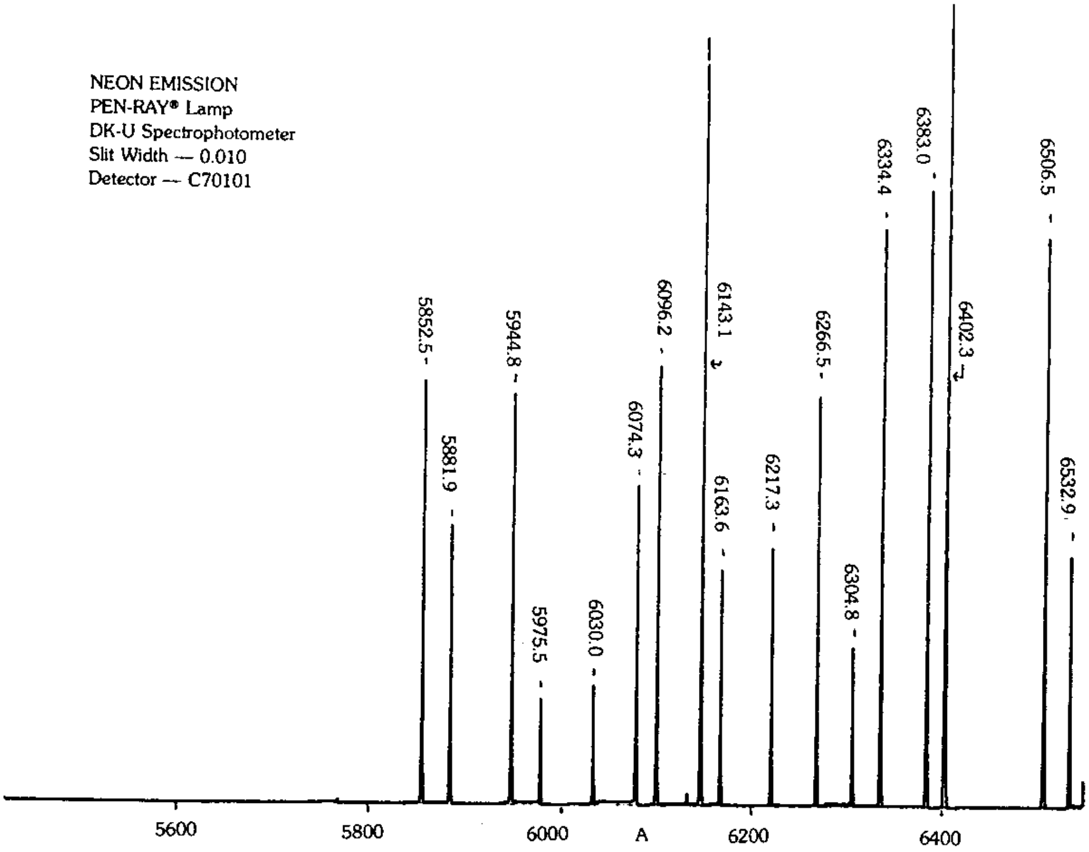
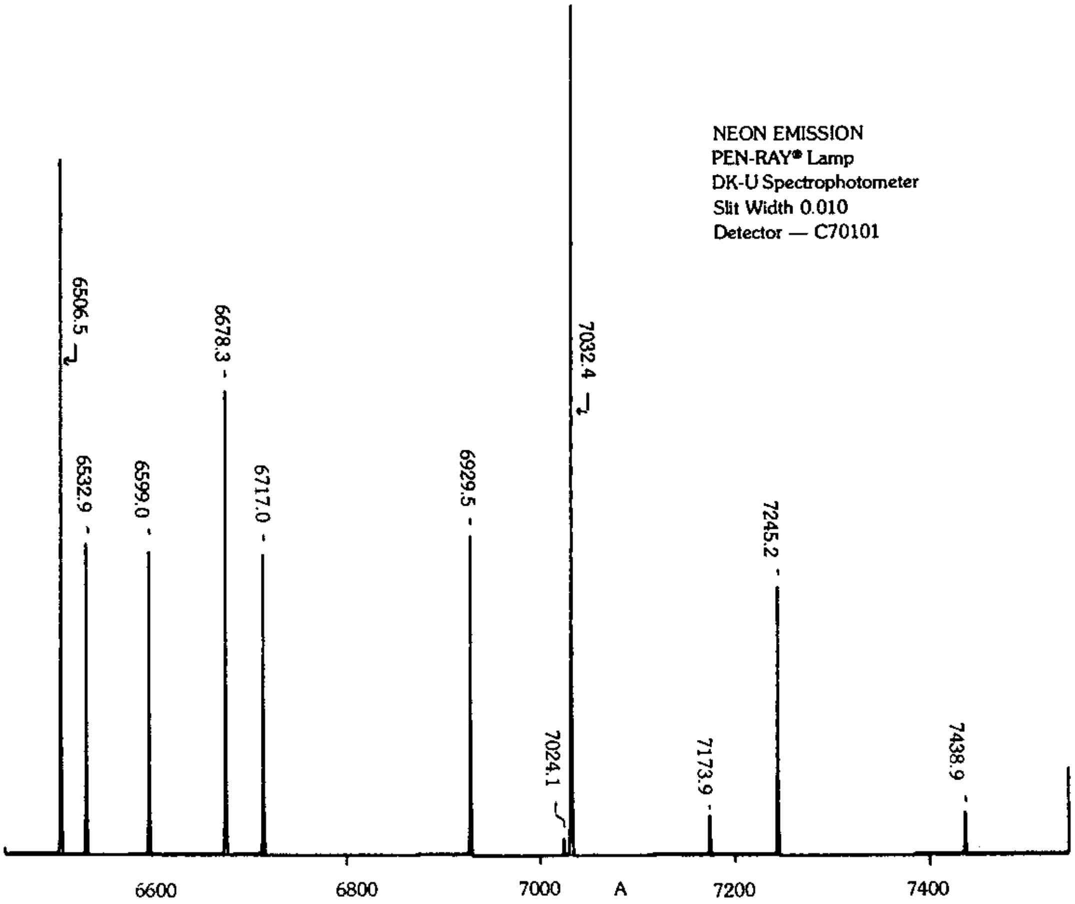
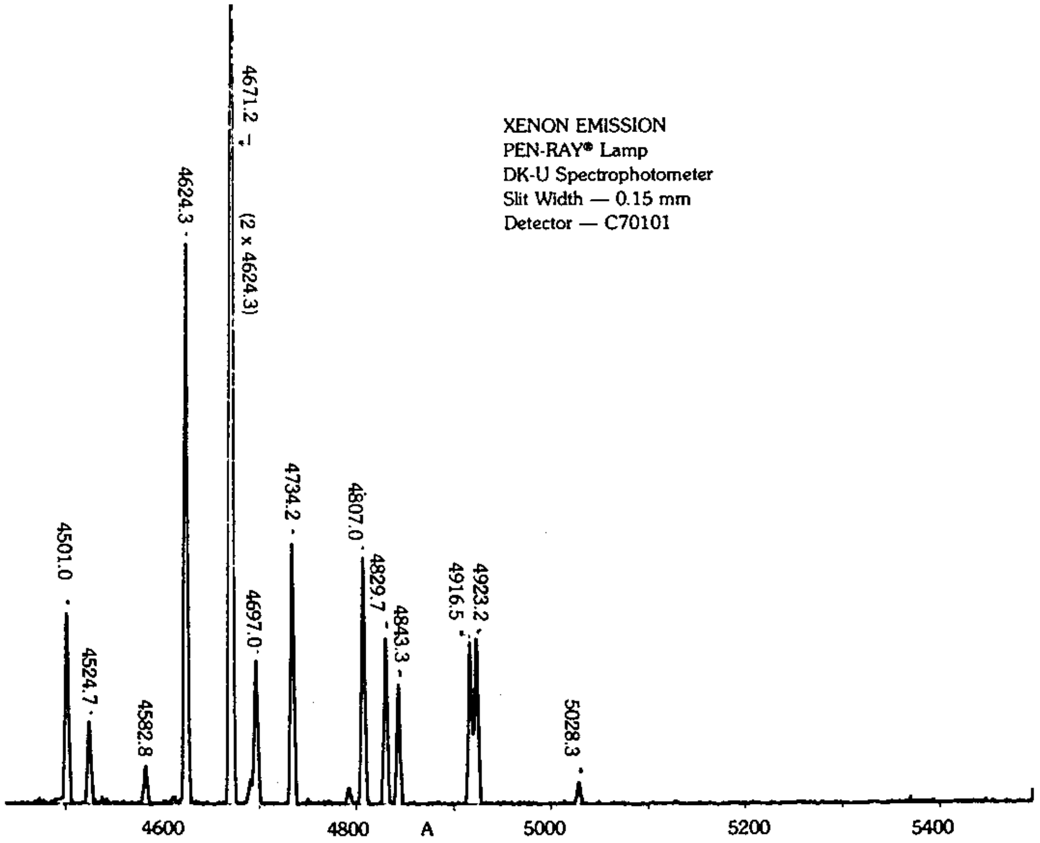


FIGURE 26



NEON EMISSION
PEN-RAY® Lamp
DK-U Spectrophotometer
Slit Width 0.010
Detector — C70101

FIGURE 27



XENON EMISSION
PEN-RAY® Lamp
DK-U Spectrophotometer
Slit Width — 0.15 mm
Detector — C70101

FIGURE 28

XENON EMISSION
PEN-RAY® Lamp
DK-U Spectrophotometer
Slit Width 0.15 mm
Detector — C70101

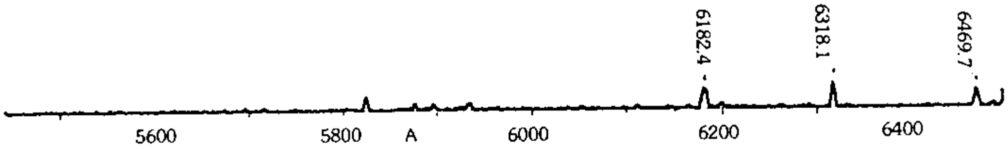


FIGURE 29

XENON EMISSION
PEN-RAY® Lamp
DK-U Spectrophotometer
Slit Width — 0.15 mm
Detector — C70101

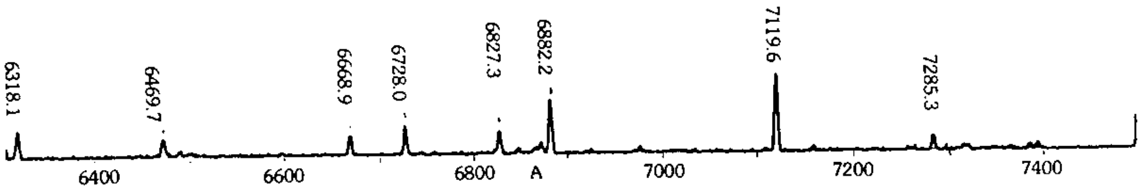


FIGURE 30

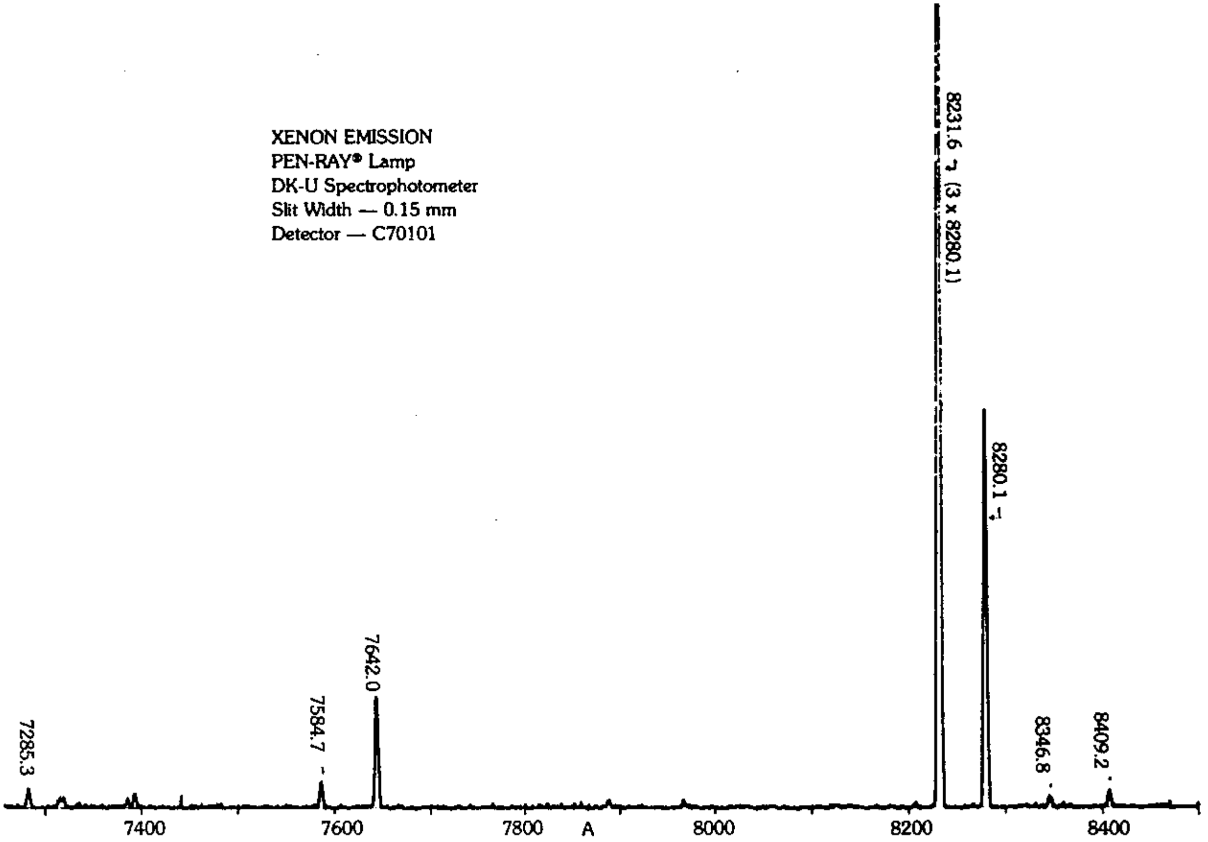


FIGURE 31

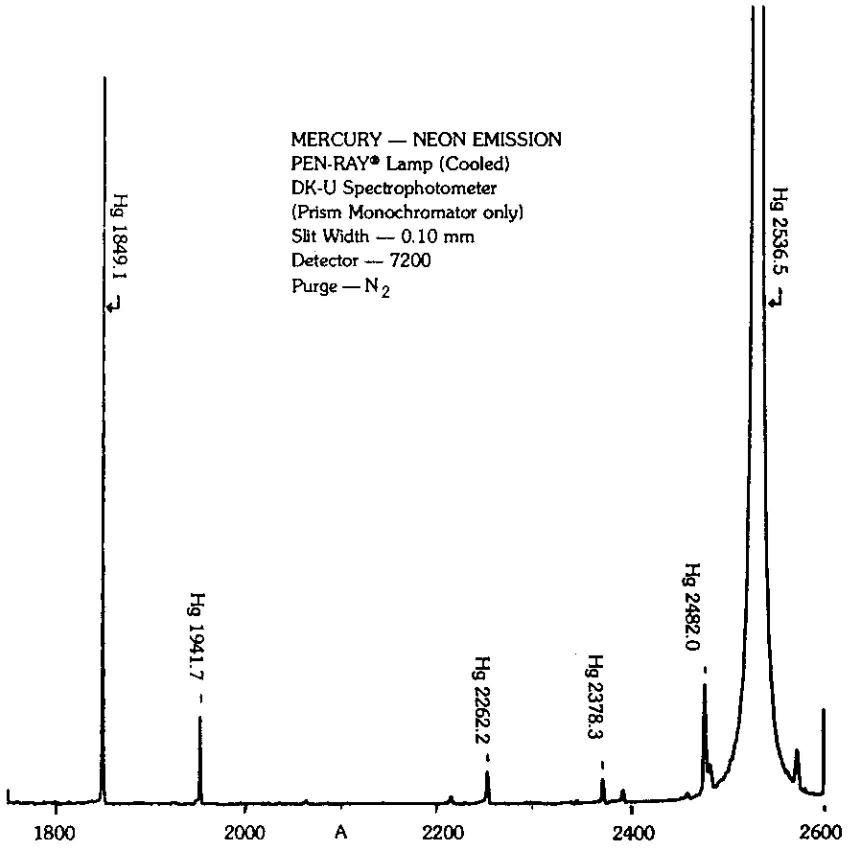


FIGURE 32

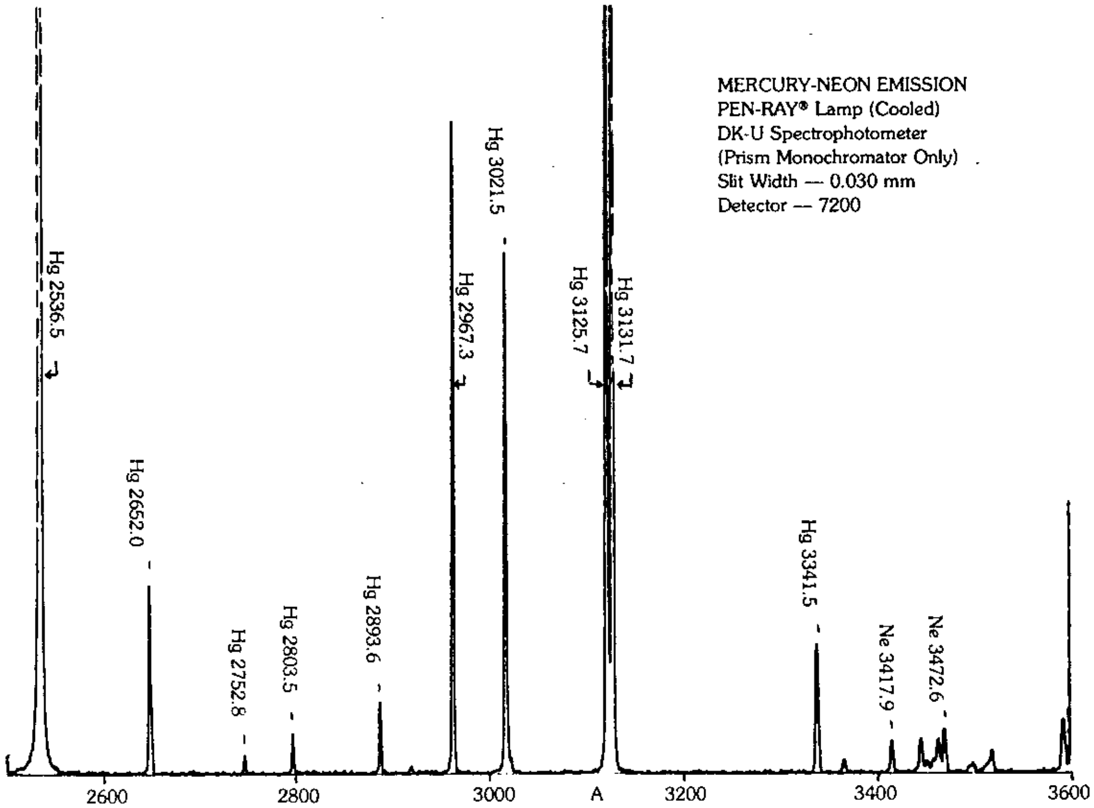


FIGURE 33

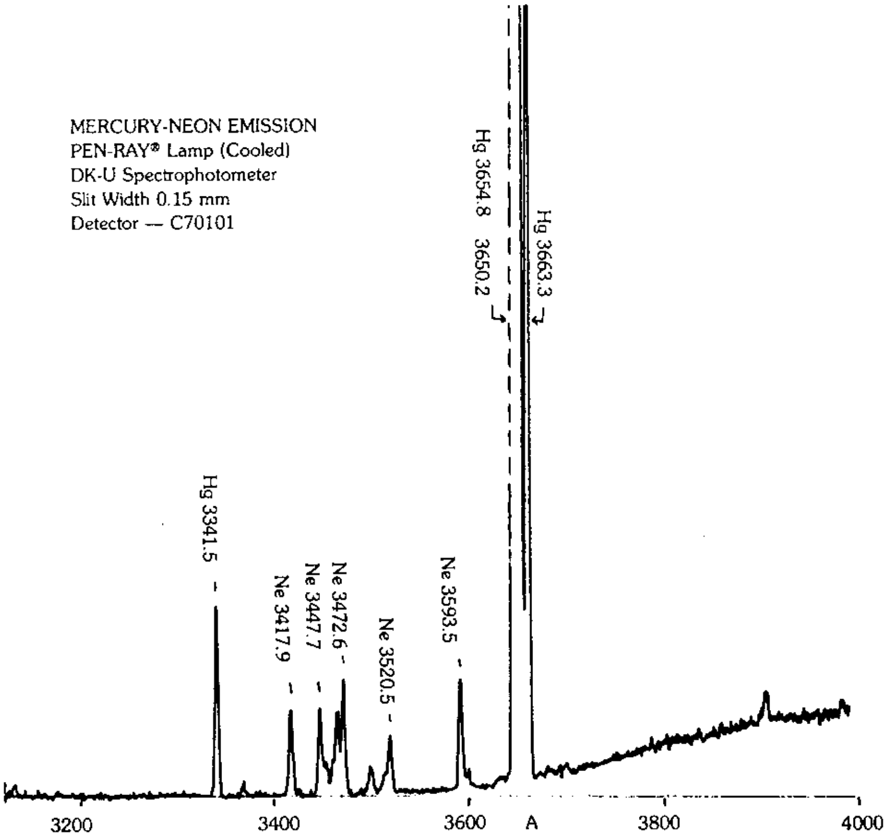


FIGURE 34

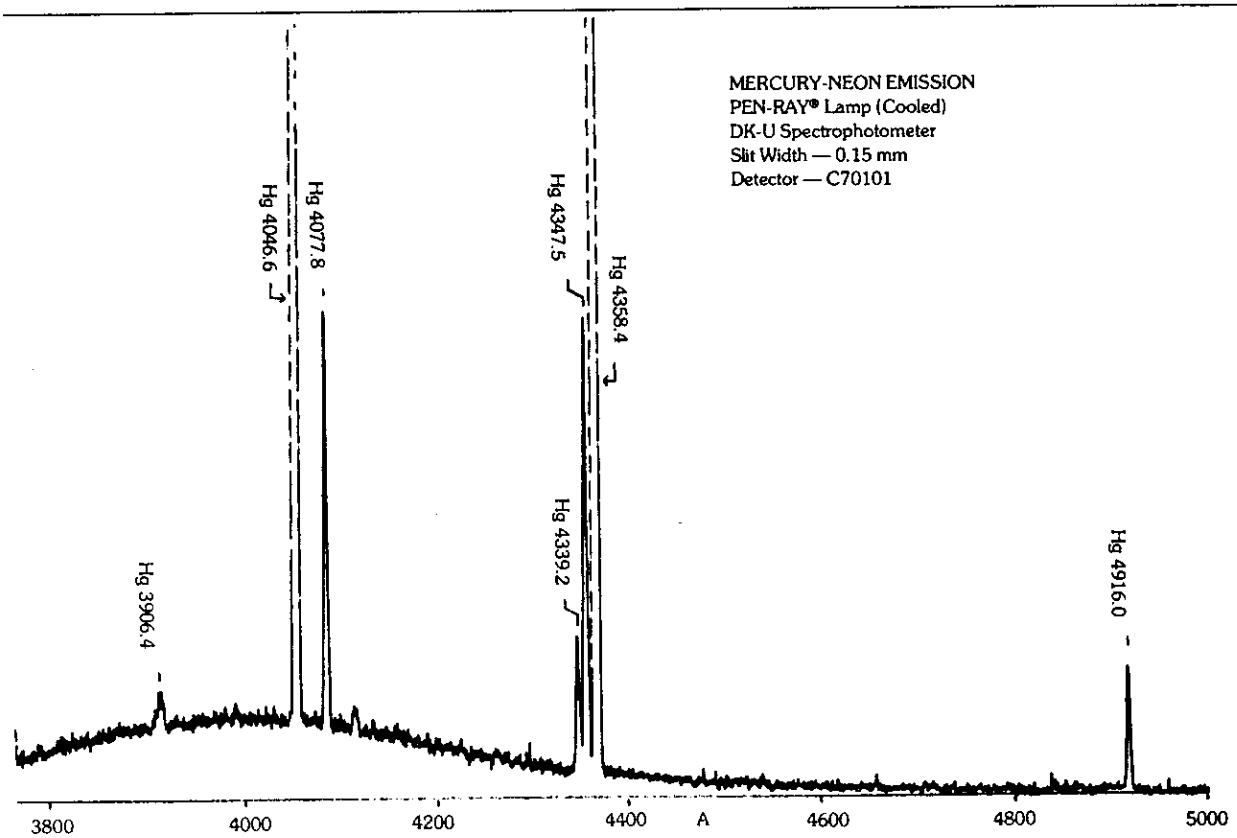


FIGURE 35

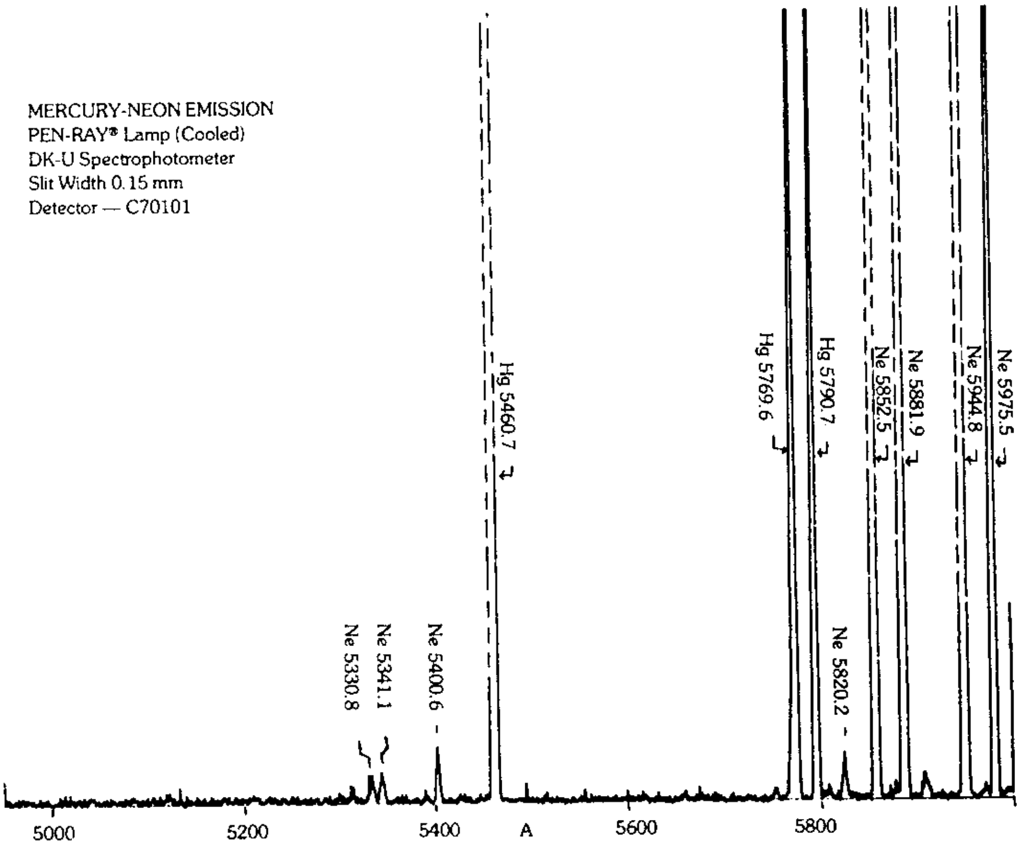


FIGURE 36

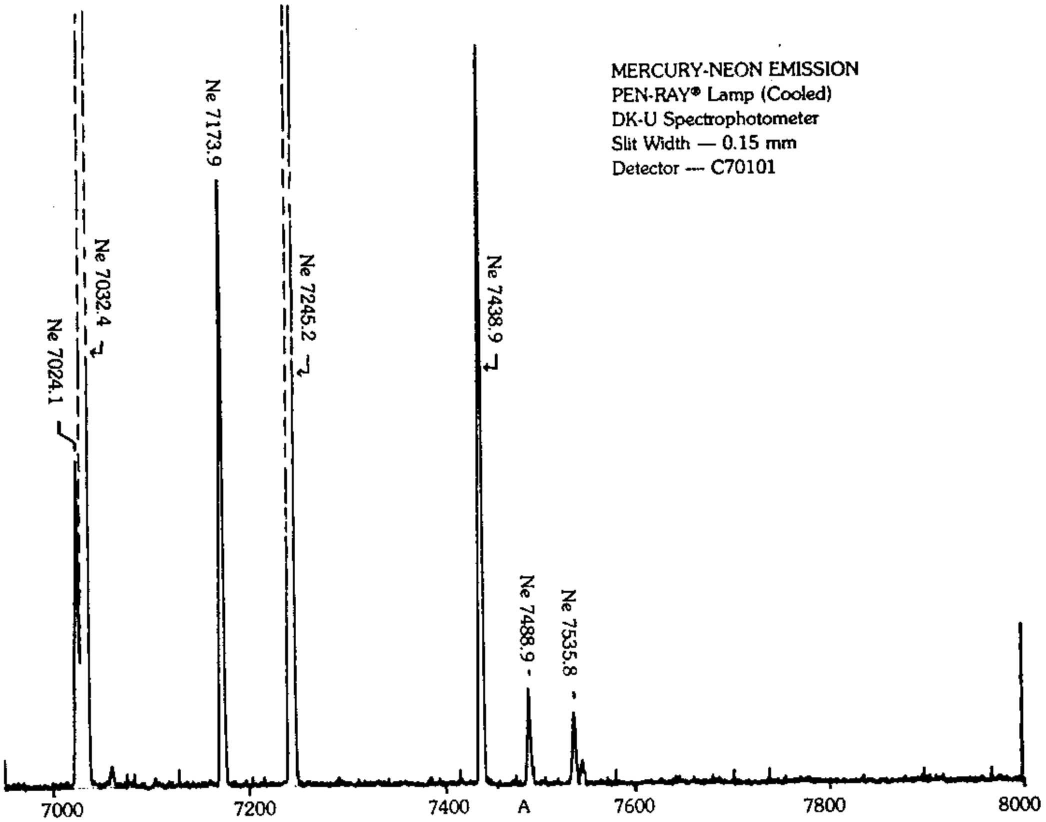


FIGURE 37

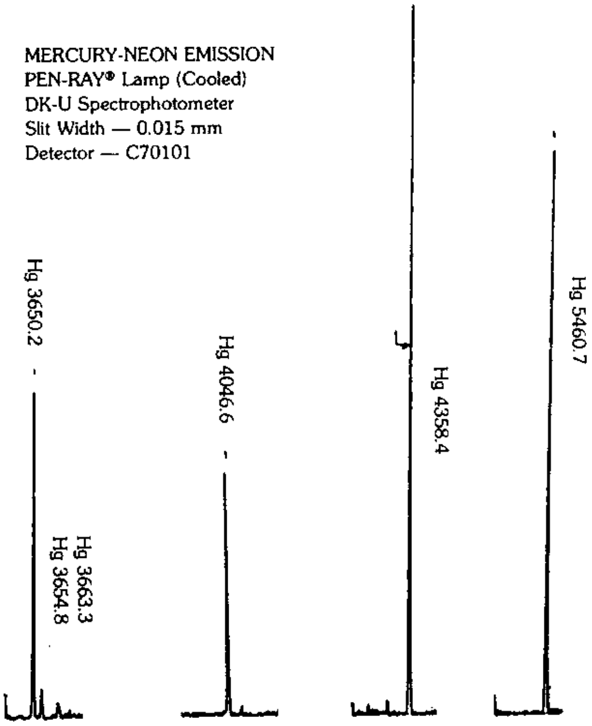


FIGURE 38

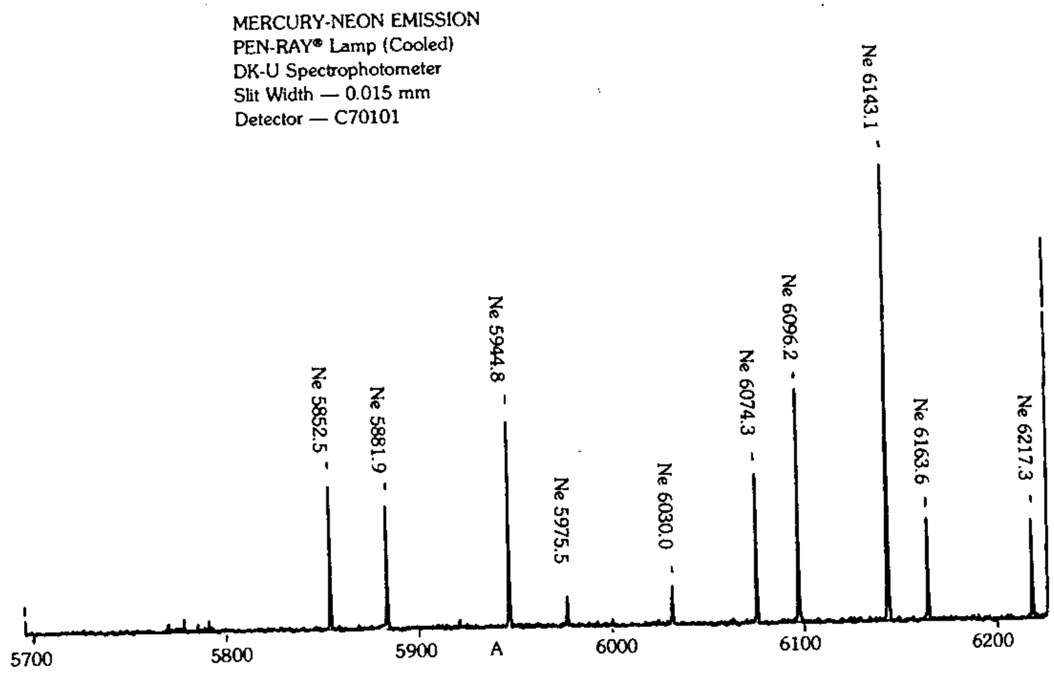


FIGURE 39

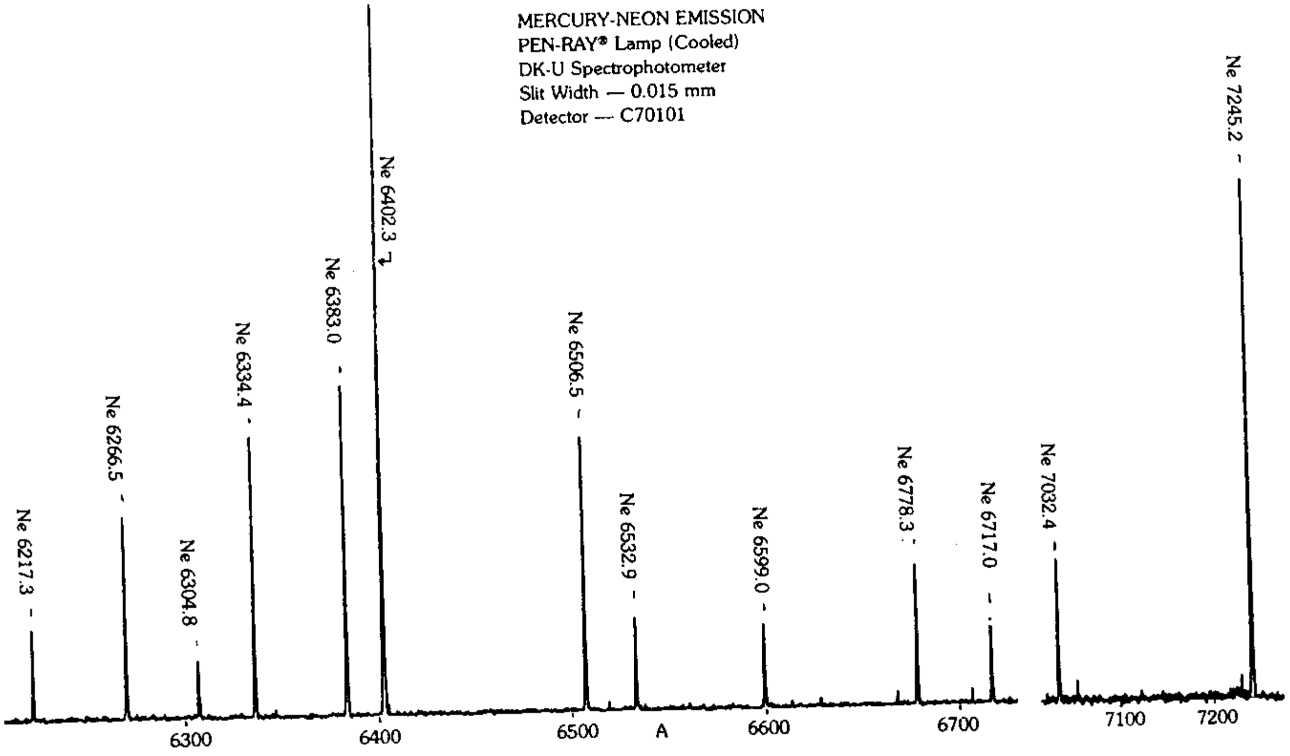


FIGURE 40

UVP
ULTRA-VIOLET PRODUCTS

http://www.uvp.com

Corporate Headquarters: UVP, Inc. 2066 W. 11th Street, Upland, CA 91786 USA
Telephone: (800)452-6788 or (909)946-3197 **Fax:** (909)946-3597

European Sales Operations: Ultra-Violet Products Limited, Science Park, Milton Road,
Cambridge CB4 4FH England **Telephone:** 44(0)1223-420022 **Fax:** 44(0)1223-420561

6 The Franck-Hertz experiment

Abstract: In this lab we will carry out the Franck-Hertz experiment, which first observed quantized energy exchange in electron-atom scattering. We'll use this to measure the first excitation energy of Mercury.

6.1 Pre-lab preparation

In the atomic spectra lab, we saw that atoms emit photons of specific, quantized energies. In this lab, we'll see that they also absorb energy in quantized steps from an electron's kinetic energy. This experiment was first done in 1914 by J. Franck and G. Hertz, and it helped to validate the quantum theory.

The experiment generates a current of electrons from a cathode and accelerates them with a voltage on the anode, V_a, so the maximum kinetic energy that can be obtained by the electrons is eV_a. A low pressure gas of mercury atoms in the region between the cathode and anode causes the moving electrons to scatter off the mercury atoms. If the atoms could absorb any energy from the electrons, we'd expect that the electrons would lose energy even when they are moving slowly. But, if the atoms will only absorb discrete energies from the electrons, then we'd expect to see a sudden onset of energy transfer when the electron kinetic energy becomes equal to the lowest (or first) excitation energy of the mercury atoms. We can measure this by slowly increasing V_a and measuring the kinetic energy of the electrons that reach the anode. A sudden drop should occur when eV_a reaches the first excitation energy eV_1.

If the voltage is increased above V_a, the electron can excite a mercury atom in one collision and then regain energy afterward. When V_a reaches $2V_1$, an electron can undergo two interactions, and be left with little energy when it reaches the anode. This continues as the voltage is increased further; we should see drops in the final kinetic energy each time V_a crosses an integer multiple of mercury's excitation energy.

There are two complications in the actual experiment. First, it is not easy, or necessary, to directly measure the final kinetic energy of the electrons. It is simpler, and sufficient, to just measure how many electrons are above some kinetic energy threshold. We can do that by using a grid for the anode, so the electrons can pass through it, and then place a third electrode beyond it. Applying a retarding voltage, $V_r > V_a$, to that electrode will allow only electrons with kinetic energy above $e(V_r - V_a)$ to reach it. Measuring the current collected by that electrode will measure how many electrons have sufficient kinetic energy to reach it. We can detect when there is a drop in electron kinetic energy due to excitation of the mercury atoms be observing a drop in the current through the "collection electrode".

The second complication is that there are small offsets to the kinetic energy of the electrons due to the contact potential of the electrodes. The thermionic emission of electrons from the cathode requires a small amount of energy, similar to the work function of the metal that we encountered in measurements of the photoelectric effect. The voltage when the first drop occurs is not a good measure of V_1 due to bias from these offsets. However, measuring the *difference in voltage between subsequent current drops* removes the bias and can be used to determine V_1.

6.1.1 Apparatus

A diagram and photo of the apparatus you'll use is shown in Figure 21. In the diagram, the gas volume is shown by the elliptical object. The cathode, and its heater, is shown at the bottom of the volume. The anode grid is represented by the dashed line, and above that is the collection electrode. The gas volume is housed within an oven that is heated by the power supply at the upper left, and its power is controlled by the large variac. To produce thermionic emission, the cathode is heated by a filament controlled by the smaller variac, shown at the bottom left. The anode voltage, V_a, is set by the power supply, and

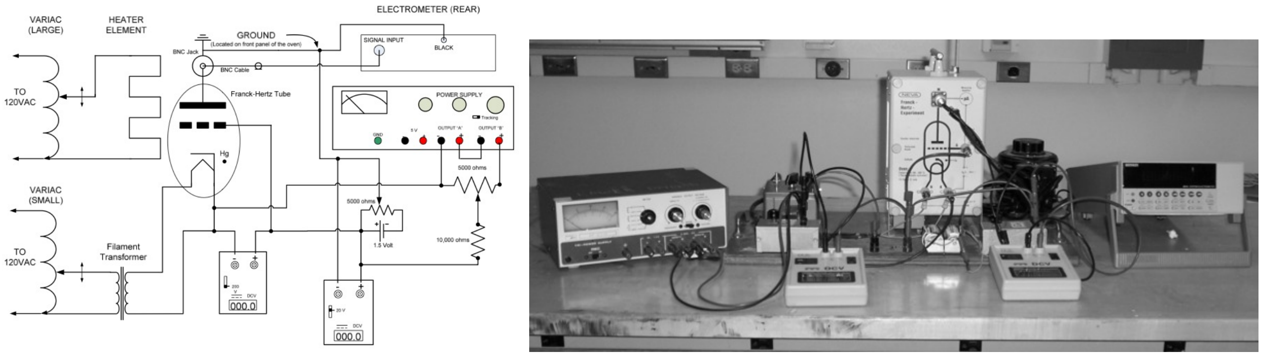


Figure 21: (a) Diagram of the apparatus. (b) Photo of the apparatus.

adjusted with a variable resistor. The retarding voltage on the collection electrode, V_r, is set by a 1.5 V battery with a variable resistor. Each of the voltages, V_a and V_r, are measured by voltmeters. The current reaching the collection electrode is measured by the electrometer, which is the same device we used in the photoelectric effect lab.

6.1.2 Safety and care of equipment

The potential hazards in this lab involve the high temperature of the oven and the moderately high voltage of the power supply. Both hazards are mitigated by being contained within enclosures.

6.2 Getting started and gaining familiarity

The first thing to do is insert the thermometer into the oven so that its bulb is even with the mid section of the tube. Then turn the oven on. The target temperature is in the $170 - 190^\circ$ C range. It can be adjusted with the large variac knob. You should set it initially to about "50" on the variac's scale, but monitor the temperature frequently and adjust accordingly. Note that the thermal response time is large. It may take 10 - 15 minutes to warm up, and adjusting the variac knob will not cause an immediate change. It is best to make small changes and monitor frequently. (A previous lab session may have already heated the oven for you).

Check the setup and confirm that it is wired as shown in Figure 21(a). Ask the TA for assistance if anything does not make sense.

Turn on the electrometer and correct any zero offsets. The procedure for this is the same as in the photoelectric effect lab, and is listed in Figure 22.

When the oven temperature has stabilized at about 180° C, set the retarding potential to 1.3 V and then turn on the power supply. Be sure that it is set to "Tracking mode" and the output voltage is set to 40 V (far right knob, or "master control", depending on the power supply model). Then adjust the $5\text{k}\Omega$ potentiometer that controls V_a up till you measure about 40 V on V_a's voltmeter.

Set the knob on the small variac to zero and then turn on the variac. It controls the cathode's filament heater which causes electrons to be thermionically emitted from the cathode. Slowly increase the variac knob until you get a current reading on the electrometer of between 0.75 and 1.0 nA, but don't go beyond that current. **Caution:** if you exceed 1 nA before the variac has been turned more than about halfway, shut it off immediately and notify the TA.

Now that everything is on and functioning, turn V_a down to zero so you can begin making measurements.

Zero Correction procedure to remove offsets

- Enable Zero Check mode. Press ZCHK; you should see a “ZC” message appear at the right side of the display.
- Select the current measuring mode by pressing I (if not already in that mode).
- Select the appropriate measurement range.
- Perform a zero correction by pressing ZCOR; you should see a “ZZ” message appear in the display.
- Press ZCHK to disable the Zero Check mode; you should now see a “CZ” message appear in the display to indicate that the subsequent readings are corrected for any zero point offsets.

Figure 22: Zero correction procedure.

6.3 Measuring the excitation energy

Slowly increase V_a and you should see the electrometer current, I, increase and then decrease after it crosses the point where electrons can excite mercury atoms. Adjust V_a so that it is right at the point where the current peaks, and record both V_a and the current. Then continue to increase V_a and you should see that the current goes through a minimum. Record I and V_a at that minimum.

Continue to increase V_a and similarly record the voltage and current for each of the subsequent peaks and minima until you reach the maximum possible value of $V_a = 40$ V. You should see about six cycles of current peaks and minima.

In addition to recording the position of the peaks and minima, you should record I vs V_a regularly while increasing V_a. This will allow you to make a plot of the full $I(V)$ curve, which illustrates the behavior.

From your measurements of V_a at the peaks, determine the excitation energy of mercury. Keep in mind that the first peak is biased by the contact potential offsets, so you should use the differences between the subsequent peaks for your measurement.

6.4 Going Beyond

While you can obtain a measurement from a single scan, or even from just a couple peak positions, it is always a good idea to measure things repeatedly. The variation between multiple measurements helps to validate the results. You’d be wise to take more than one scan.

Any additional time available when you complete this lab could be used to further your analysis of the measurements from your personal experiment.