Measurement and Measurement Error

PHYS 1301 F98
Prof. T.E. Coan
Last edit: 5 Aug '98

<u>Introduction</u>

Physics makes very general yet quite detailed statements about how the physical universe works. These statements are organized or grouped together in such a way that they provide a model or a kind of coherent picture about **how and why** the universe works the way it does. These sets of statements are called "theories" and are much more than a simple list of "facts and figures" like you might find in an almanac or a telephone book. (Even though almanacs and telephone books are quite useful!) A good physics theory is far more interested in principles than simple "facts". Noting that the moon appears regularly in the night sky is far less interesting than understanding **why** it does so.

We have confidence that a particular physics theory is telling us something interesting about the physical universe because we are able to test **quantitatively** its predictions or statements about the universe. In fact, all physics (and scientific) theories have this "put up or shut up" quality to them. For something to be called a physics "theory " in the first place, it **must** make sufficiently quantitative statements about the universe that can be then quantitatively tested. These tests are called "experiments". The statement, "My girlfriend is the most charming woman in the world," however true it may be, has no business being in a physics theory because it simply cannot be quantitatively tested. If the experimental measurements are performed correctly and the observations are inconsistent with the theory, then at least some portion of the theory is wrong and needs to be modified. On the other hand, consistency between the theoretical predictions and the experimental observations never decisively proves a theory true. It is always possible that a more comprehensive theory can be developed which will also be consistent with these same experimental observations. Physicists, as well as other scientists, give a kind of provisional assent to their theories, however successfully they might agree with experimental observation. Comfortably working with uncertainty is a professional necessity for them.

It is clear then that making careful quantitative measurements are important if we want to claim that a particular physics theory explains something about the world around us. Measurements represent some physical quantity. For example, 2.1 meter represents a distance, 7 kilograms represents a mass and 9.3 seconds represents a time. Notice that each of these quantities has a number like 9.3 and a unit, "seconds". The number tells you the amount and the unit tells you the thing that you are talking about, in this case seconds. **Both the number and the unit** are required to specify a measured quantity.

Precision

There is a certain inherent inaccuracy or variation in the measurements we make in the laboratory. This inherent inaccuracy or variation is called experimental "error" and in no way is meant to imply incompetence on the part of the experimenter. It merely reflects the fact that our measuring instruments are not perfect. This lack of perfection in our measuring procedure is to be contrasted with mistakes like adding two numbers incorrectly or incorrectly writing down a number from an instrument. Those mistakes have nothing to do with what I mean by experimental error and everything to do with the competence of the experimenter! Understanding and quantifying measurement error is important in experimental science because it is a measure of how seriously we should believe (or not believe) our theories abut how the world works.

When we make measurements in the laboratory, we should distinguish between the **precision** and the **accuracy** of these measurements. If I measure my mass to be 120.317 kilograms, that is a very *precise* measurement because it is very specific. It also happens to be a very *inaccurate* measurement because I am not quite that fat. My mass is considerably less, something like 85 kilograms. So, when we say that we have made a precise measurement we can also say that we have made a very specific measurement. When we say we have made an accurate measurement we can also say that we have made a correct measurement.

The so-called number of significant figures that it contains indicates the precision of a measurement. In our mass example, the quantity 120.317 kilograms has 6 significant figures. This is rather precise as measurements go and is considerably more precise than anything you will measure in this course. Again, the fact that a measurement is precise does not make it accurate, just specific. In any event, it is useful to be familiar with how to recognize the number of significant figures in a number. The technique is this:

1. The leftmost non-zero digit is the most significant digit;
2. If there is no decimal point, the rightmost non-zero digit is the least significant;
3. If there is a decimal point, the rightmost digit is the least significant, even if it is zero.
4. All digits between the least significant and the most significant (inclusive) are themselves significant.

For example, the following four numbers have 4 significant figures:
2,314   2,314,000   2.314   9009   9.009   0.000009009   9.000

No physical measurement is completely exact or even precise. The apparatus or the skill of the observer always limits accuracy and precision. For example, often physical measurements are made by reading a scale of some sort (ruler, thermometer, dial gauge, etc.). The fineness of the scale markings is limited and the width of the scale lines is greater than zero. In **every** case the final figure of the reading must be estimated and is therefore somewhat inaccurate. This last figure does indeed contain some useful information about the measured quantity, uncertain as it is. A significant figure is one that is reasonably trustworthy. **One and only one** estimated or uncertain

figure is retained and used in the measurement. This is the least significant figure. It is also important to understand approximately the magnitude of the uncertainty so that you can state it with the measurement.

For example, a length measurement of 2.500 meters (4 significant figures) is one digit less uncertain that a measurement of 2.50 meters (3 significant figures). If your measuring instrument is capable of 4 digit precision then you can distinguish between an item 2.500 meters long and one 2.501 meters long. If your instrument were precise to only 3 significant figures, then both items would appear to be 2.50 meters long and you could not distinguish between them. Again, infinite precision is not possible. All instruments have limits. It is important to know how much precision your instrument is capable of. For example, it would be silly to measure something with an inexpensive ruler and then announce that the object was 9.3756283 centimeters long. With such a ruler, the finest scale markings are typically 1 millimeter (0.1 cm) apart. You can estimate to 0.5 millimeter, possibly 0.3 millimeter at the best. An estimate of 9.38 centimeters with an uncertainty of +/- .03 centimeters would be more appropriate.

Error

In general, two types of error limit precision and accuracy: **random errors** and **systematic errors**. What could cause several measurements of the same thing to yield slightly different values? This is an example of random error. This means that the *variations* in your measurements have no fixed or predictable pattern. Your measured value could be a bit high or a bit low of the correct value. In fact, with random errors an individual measurement of an object is just as likely to be a bit too high as to be a bit too low. For example, using a ruler often produces random errors because you will not consistently get the ruler's zero line **exactly** aligned with the edge of the measured object every time you make a measurement.

The other general kind of error is a systematic error. These types of errors are often insidious. Measurements may display good precision yet be very inaccurate. Often, measurements afflicted with systematic errors are usually *consistently* too high or too low from the accurate value. The measurements are "systematically" wrong. For example, weighing yourself with a scale that has a bad spring inside of it will yield weight measurements that are inaccurate, either too high or too low, even though the weight measurements may seem reasonable. For example, you might notice that you seem to gain weight after Thanksgiving holiday and lose weight after being ill. These observations may in fact be true but the actual value for your weight will be inaccurate. It is often very difficult to catch systematic errors. To do so requires that you understand very carefully the workings of your measuring devices.

We can estimate accuracy and precision for a group of measurements and we can find the absolute and relative uncertainties by simple calculation. The absolute uncertainty, in percent, is given by:

$$\%A = |\ (\ (m\text{-}V)\ /\ V)\ |\ \text{x}\ 100\%,$$

where *m* is a measured value and *V* is the correct value. Assuming that you somehow know the correct value, this uncertainty is computed after all measurements have been made. It will tell you how different your measurement is, in percent, from the actual value of the quantity you are trying to measure.

Since we don't normally know the true value of the quantity we are measuring, (otherwise why would we need to measure it?), we need some way to estimate how reliable our measurements are. That is, we need some measure of how *precise* our measurements are. If someone asks us to measure the distance between the Earth and the Sun, they will also be interested in knowing how precisely we think we know our answer. We need to be able to give them an answer in a compact form that indicates our best estimate and our uncertainty in that estimate: $x = x_{best} \pm dx$. Estimating the uncertainty or error – which is a measure of the precision of our measurement – is usually the hard part and what we concentrate on now.

One way of estimating the precision of our measurements of some quantity is to measure the quantity many times and then use the data to estimate how precise our measurement is. The procedure is tedious to describe in words but very easy if you use symbols. We'll use words first to show you the advantage of using symbols. After making multiple measurements m_i, where the subscript *i* labels a particular measurement of the quantity (*i* = 1 if we are talking about the first measurement or *i* = 5 if we are talking about the 5th measurement), we form the average $\overline{m}$ of all the measurements. Next, we a sum up the square of the difference between $\overline{m}$ and all the m_i's. Continuing, we divide this sum by *N* – 1, where *N* is the total number of measurements we made. Finally, we take the square root of this number to find the number we really care about. Summarizing, and this time using only symbols, we have

$$s^2 = \frac{\sum_i (m_i - \overline{m})^2}{N - 1}.$$

The symbol s^2 is called the "variance of the mean" and its square root σ (the "standard deviation of the mean") is the number we are interested in. The value of σ is what we would substitute for δ*x* above. It summarizes the variation or "spread" in our data when we make multiple measurements of the same quantity. It is a measure of the precision or uncertainty of our final answer x_{best}. Large values of σ mean our individual measurements varied substantially as we made them over and over again. In such a case we do not expect that our final answer will be particularly precise. Conversely, if σ is small, our individual measurements varied little. This is easy to see by looking at the numerator for our formula for s^2. The formula tells us to subtract the values of our individual measurements from the average for all the measurements and then add all these differences. So, if all our measurements are nearly the same, then the average for the measurements will not be very different from any individual measurement. The numerical value of the numerator will be small and so will s^2. Notice also that we divide the numerator by *N* –*1* so that the entire fraction is a kind of

average deviation of the individual measurements from the mean of the measurements. From this you can see why σ is baptized "the standard deviation of the mean."

Although a bit tedious, the above procedure for estimating the uncertainty σ in our final answer is straightforward. But what do you do if you know or can estimate the uncertainties in *individual* measurements? For example, suppose you can measure lengths to a precision of +/- 3 mm. You measure the length of some quantity many times and now you want to make a statement about the uncertainty in your value of x_{best}. You might think that you ought to be able to make use of this knowledge of the uncertainty of an individual measurement. You are correct. To get the uncertainty in your <u>final</u> answer, you need to add the uncertainties in the individual measurements. But, you need to add them in a special way. Specifically, you compute the sum

$$\frac{1}{s^2} = \sum \frac{1}{s_i^2},$$

where the s_i are the uncertainties in the *individual* measurements. In words, this formula says to add the squares of the reciprocals of the individual uncertainties. This may seem like a perverse way to proceed but it is actually quite sensible. If a particular measurement has a very large uncertainty (because you forgot to turn on part of the measuring device), then you do not expect this particular measurement to be very useful for finding your final answer. Well, the above formula tells you this. If you take the reciprocal of a large number then you will get a small number. Adding a small number to some other number doesn't always change the original number very much so in some cases you can just neglect the small number.

Remember, what you really care about ultimately is the uncertainty in your final answer and that is given by σ. Either one of our computational recipes is appropriate.

<u>Odds and Ends</u>

Taking the mean (average) value for a set of measurements is a good way of reducing the effects of random errors and is the preferred way to get your value for x_{best}. As you increase the number of measurements, the mean value of these measurements will more closely resemble the actual value of the quantity you are trying to measure, assuming there are no systematic errors present. This is because you are just as likely to measure a value that is slightly too high as one that is too low, so that the random errors will "average themselves out". The magnitude of the relative errors for the individual measurements gives you an idea of the uncertainty in your value.

Averaging your measurements will not help you when there are systematic errors present. In fact, there is no standard way to deal with systematic measurements, which is why they are so troublesome.

There is a final a aspect of measurement which should be aware of. We should be very specific about exactly what it is we are trying to measure. For example, in our weight example above, does it mean with or without clothes? At what time of day? Before or after a meal? Consider the question, "How far is it to Fort Worth?" From

where in Dallas to where in Ft. Worth? Straight line distance or road distance? You should be very clear about the meaning of the quantity you are trying to measure so that people can accurately interpret the significance of your results.

Summary Questions

- What is the difference between accuracy and precision and why is this difference significant?
- How many significant figures are in 13600? 0.12040? 00045?
- What distinguishes a random error from a systematic error?
- What is the point of making multiple measurements of the same quantity before making a final statement about the numerical value of that quantity?

Objectives

1. To see how measurements and error analysis are a fundamental part of experimental science.

2. To make some actual measurements and analyze the errors in them.

3. To observe and understand the difference between accuracy and precision.

4. To understand the nature of random and systematic errors.

Equipment

Candle, ruler, special ruler, aluminum dish, metal rods, material to be weighed (dry ice and alcohol) and balance scale.

Procedure

1. Measure the length of the metal rods.
Using **only** the special ruler, measure the length of each piece 5 times. When you are finished, average the values to get a better measure of the piece's true length. Next, use your plastic ruler to measure your pieces again. Measure 5 times as before and compute the average to refine your measured value. Make an "eyeball" estimate of your uncertainties.

2. Measure the height of the candle flame.
Light your candle and let the flame burn steadily for a minute or so. Use the plastic ruler to measure the height of the flame. Make 10 measurements and try not to melt the ruler. Hold the ruler a small distance away from the flame Record your measurements. Estimate for each measurement its s_i. Identify the technique you used

to find the uncertainty σ of your final answer. Remember to write only a sensible number of significant figures.

3. Measure your reaction time.

Your reaction time is the time that passes between some external stimulus and your first action. We will use an old method to measure your reaction time. A falling ruler will suffice. This is what to do.

a. Have your partner hold the regular ruler vertically, holding it by the top and having the zero point toward the bottom.
b. Place your thumb and forefinger at the ruler's bottom, surrounding the zero point. Be prepared to pinch the ruler as if it were to fall. Rest your forearm on the lab table to steady your hand.
c. Your partner will drop the ruler without warning.
d. Pinch and grab the falling ruler as fast as you can. Record the distance the ruler fell. This will tell you your reaction time.
e. Compute your reaction time using Galileo's formula: $s = \dfrac{1}{2}gt^2$. The meaning of the symbols will be explained in lab.

Make 5 measurements and record the corresponding reaction times. Record **your** reaction times on **your** data sheet. Do not mix your times with your partners. This means you will make 5 measurements per person. Compute and record the individual s_i's for each measurement. Identify the technique you used to find the uncertainty σ of your final answer. Be sure that both you **and** your lab partner have your reaction times measured.

4. Measure the diameter of a closed curve.

Measure the diameter of the large closed curve. Measure this curve diameter across 6 *different* diameters of the curve. Record your measurements. Compute the average diameter. Estimate for each measurement its s_i. Identify the technique you used to find the uncertainty σ of your final answer. Remember to write only a sensible number of significant figures.

5. Measure the mass of a cold material.

Go to the instructor's table with your partner, where you will be given a cup containing some alcohol and some crushed dry ice. Using the balance on the instructor's table, measure and record the mass of the cup 6 times, at 1 minute intervals. **Warning: The dry ice and alcohol mixture is quite cold. If you stick your fingers in the mixture you will feel much pain.**

Measurement and Measurement Error

PHYS 1301 F98

Name: _____   **Section**: _____

Abstract:

Analysis:

1. Rod lengths

 Rod 1   Special Ruler   Plastic Ruler

Measurement 1   _____   _____

Measurement 2   _____   _____

Measurement 3   _____   _____

Measurement 4   _____   _____

Measurement 5   _____   _____

Measurement 6   _____   _____

Average Value   _____   _____

Uncertainty   _____   _____

Rod 2	Special Ruler	Plastic Ruler
Measurement 1	_____	_____
Measurement 2	_____	_____
Measurement 3	_____	_____
Measurement 4	_____	_____
Measurement 5	_____	_____
Measurement 6	_____	_____
Average Value	_____	_____
Uncertainty	_____	_____

a. What are the possible sources of error in this measurement?

b. How well did your measurements with the special ruler agree with those done with the plastic ruler? If there was a disagreement, what kind of error was it? Random or systematic? What caused this error?

2. Candle Flame

	Flame height	S_i
Measurement 1	_____	_____
Measurement 2	_____	_____
Measurement 3	_____	_____
Measurement 4	_____	_____
Measurement 5	_____	_____
Measurement 6	_____	_____
Measurement 7	_____	_____
Measurement 8	_____	_____
Measurement 9	_____	_____
Measurement 10	_____	_____
Average Value	_____	_____
Uncertainty	_____	_____

a. What are the possible sources of error in this measurement?

b. What might you do to get a better measurement of the flame's height?

3. Your reaction time

	Distance	Time	S_i (time)
Measurement 1	_____	_____	_____
Measurement 2	_____	_____	_____
Measurement 3	_____	_____	_____
Measurement 4	_____	_____	_____
Measurement 5	_____	_____	_____
Average value	_____	_____	
Uncertainty	_____	_____	_____

a. What are the possible sources of error in this measurement?

4. Closed Curve Diameter

	Curve diameter	s_i
Measurement 1	_____	_____
Measurement 2	_____	_____
Measurement 3	_____	_____
Measurement 4	_____	_____
Measurement 5	_____	_____
Measurement 6	_____	_____
Average value	_____	_____
Uncertainty	_____	

a. What are possible sources of error in this measurement?

b. What do the measurements tell you about the curve's diameter?

5. Mass of cold material.

Measurement 1   _____

Measurement 2   _____

Measurement 3   _____

Measurement 4   _____

Measurement 5   _____

Measurement 6   _____

a. What are possible sources of error in this measurement?

b. Do you see any pattern in the measured masses?

6. Which of your measurements (metal rods, curve diameter, flame height, etc.) was the most uncertain? Why was it so?

7. Which of your measurements (metal rods, curve diameter, flame height, etc.) was the least uncertain? Why was this so?

8. Which measurements (metal rods, curve diameter, flame height, etc.), if any, suffered from systematic error?

Conclusions: