

Part I

1. Preliminary Description of Error Analysis
 2. How to Report and Use Uncertainties
 3. Propagation of Uncertainties
 4. Statistical Analysis of Random Uncertainties
 5. The Normal Distribution
-

Part I introduces the basic ideas of error analysis as they are needed in a typical first-year, college physics laboratory. The first two chapters describe what error analysis is, why it is important, and how it can be used in a typical laboratory report. Chapter 3 describes error propagation, whereby uncertainties in the original measurements “propagate” through calculations to cause uncertainties in the calculated final answers. Chapters 4 and 5 introduce the statistical methods with which the so-called random uncertainties can be evaluated.

Chapter I

Preliminary Description of Error Analysis

Error analysis is the study and evaluation of uncertainty in measurement. Experience has shown that no measurement, however carefully made, can be completely free of uncertainties. Because the whole structure and application of science depends on measurements, the ability to evaluate these uncertainties and keep them to a minimum is crucially important.

This first chapter describes some simple measurements that illustrate the inevitable occurrence of experimental uncertainties and show the importance of knowing how large these uncertainties are. The chapter then describes how (in some simple cases, at least) the magnitude of the experimental uncertainties can be estimated realistically, often by means of little more than plain common sense.

1.1 Errors as Uncertainties

In science, the word *error* does not carry the usual connotations of the terms *mistake* or *blunder*. Error in a scientific measurement means the inevitable uncertainty that attends all measurements. As such, errors are not mistakes; you cannot eliminate them by being very careful. The best you can hope to do is to ensure that errors are as small as reasonably possible and to have a reliable estimate of how large they are. Most textbooks introduce additional definitions of error, and these are discussed later. For now, error is used exclusively in the sense of uncertainty, and the two words are used interchangeably.

1.2 Inevitability of Uncertainty

To illustrate the inevitable occurrence of uncertainties, we have only to examine any everyday measurement carefully. Consider, for example, a carpenter who must measure the height of a doorway before installing a door. As a first rough measurement, he might simply look at the doorway and estimate its height as 210 cm. This crude “measurement” is certainly subject to uncertainty. If pressed, the carpenter might express this uncertainty by admitting that the height could be anywhere between 205 cm and 215 cm.

If he wanted a more accurate measurement, he would use a tape measure and might find the height is 211.3 cm. This measurement is certainly more precise than his original estimate, but it is obviously still subject to some uncertainty, because it is impossible for him to know the height to be exactly 211.3000 cm rather than 211.3001 cm, for example.

This remaining uncertainty has many sources, several of which are discussed in this book. Some causes could be removed if the carpenter took enough trouble. For example, one source of uncertainty might be that poor lighting hampers reading of the tape; this problem could be corrected by improving the lighting.

On the other hand, some sources of uncertainty are intrinsic to the process of measurement and can never be removed entirely. For example, let us suppose the carpenter's tape is graduated in half-centimeters. The top of the door probably will not coincide precisely with one of the half-centimeter marks, and if it does not, the carpenter must *estimate* just where the top lies between two marks. Even if the top happens to coincide with one of the marks, the mark itself is perhaps a millimeter wide; so he must estimate just where the top lies within the mark. In either case, the carpenter ultimately must estimate where the top of the door lies relative to the markings on the tape, and this necessity causes some uncertainty in the measurement.

By buying a better tape with closer and finer markings, the carpenter can reduce his uncertainty but cannot eliminate it entirely. If he becomes obsessively determined to find the height of the door with the greatest precision technically possible, he could buy an expensive laser interferometer. But even the precision of an interferometer is limited to distances of the order of the wavelength of light (about 0.5×10^{-6} meters). Although the carpenter would now be able to measure the height with fantastic precision, he still would not know the height of the doorway *exactly*.

Furthermore, as our carpenter strives for greater precision, he will encounter an important problem of principle. He will certainly find that the height is different in different places. Even in one place, he will find that the height varies if the temperature and humidity vary, or even if he accidentally rubs off a thin layer of dirt. In other words, he will find that there is no such thing as *the* height of the doorway. This kind of problem is called a *problem of definition* (the height of the door is not a well-defined quantity) and plays an important role in many scientific measurements.

Our carpenter's experiences illustrate a point generally found to be true, that is, that no physical quantity (a length, time, or temperature, for example) can be measured with complete certainty. With care, we may be able to reduce the uncertainties until they are extremely small, but to eliminate them entirely is impossible.

In everyday measurements, we do not usually bother to discuss uncertainties. Sometimes the uncertainties simply are not interesting. If we say that the distance between home and school is 3 miles, whether this means "somewhere between 2.5 and 3.5 miles" or "somewhere between 2.99 and 3.01 miles" is usually unimportant. Often the uncertainties are important but can be allowed for instinctively and without explicit consideration. When our carpenter fits his door, he must know its height with an uncertainty that is less than 1 mm or so. As long as the uncertainty is this small, the door will (for all practical purposes) be a perfect fit, and his concern with error analysis is at an end.

1.3 Importance of Knowing the Uncertainties

Our example of the carpenter measuring a doorway illustrates how uncertainties are always present in measurements. Let us now consider an example that illustrates more clearly the crucial importance of knowing how big these uncertainties are.

Suppose we are faced with a problem like the one said to have been solved by Archimedes. We are asked to find out whether a crown is made of 18-karat gold, as claimed, or a cheaper alloy. Following Archimedes, we decide to test the crown's density ρ knowing that the densities of 18-karat gold and the suspected alloy are

$$\rho_{\text{gold}} = 15.5 \text{ gram/cm}^3$$

and

$$\rho_{\text{alloy}} = 13.8 \text{ gram/cm}^3.$$

If we can measure the density of the crown, we should be able (as Archimedes suggested) to decide whether the crown is really gold by comparing ρ with the known densities ρ_{gold} and ρ_{alloy} .

Suppose we summon two experts in the measurement of density. The first expert, George, might make a quick measurement of ρ and report that his best estimate for ρ is 15 and that it almost certainly lies between 13.5 and 16.5 gram/cm³. Our second expert, Martha, might take a little longer and then report a best estimate of 13.9 and a probable range from 13.7 to 14.1 gram/cm³. The findings of our two experts are summarized in Figure 1.1.

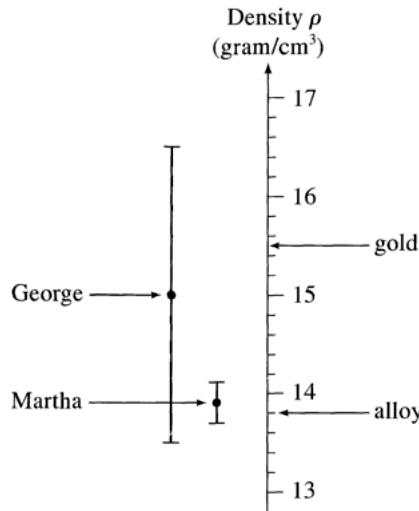


Figure 1.1. Two measurements of the density of a supposedly gold crown. The two black dots show George's and Martha's best estimates for the density; the two vertical error bars show their margins of error, the ranges within which they believe the density probably lies. George's uncertainty is so large that both gold and the suspected alloy fall within his margins of error; therefore, his measurement does not determine which metal was used. Martha's uncertainty is appreciably smaller, and her measurement shows clearly that the crown is not made of gold.

The first point to notice about these results is that although Martha's measurement is much more precise, George's measurement is probably also correct. Each expert states a range within which he or she is confident ρ lies, and these ranges overlap; so it is perfectly possible (and even probable) that both statements are correct.

Note next that the uncertainty in George's measurement is so large that his results are of no use. The densities of 18-karat gold and of the alloy both lie within his range, from 13.5 to 16.5 gram/cm³; so no conclusion can be drawn from George's measurements. On the other hand, Martha's measurements indicate clearly that the crown is not genuine; the density of the suspected alloy, 13.8, lies comfortably inside Martha's estimated range of 13.7 to 14.1, but that of 18-karat gold, 15.5, is far outside it. Evidently, if the measurements are to allow a conclusion, the experimental uncertainties must not be too large. The uncertainties do not need to be extremely small, however. In this respect, our example is typical of many scientific measurements, for which uncertainties have to be reasonably small (perhaps a few percent of the measured value) but for which extreme precision is often unnecessary.

Because our decision hinges on Martha's claim that ρ lies between 13.7 and 14.1 gram/cm³, she must give us sufficient reason to believe her claim. In other words, she must *justify* her stated range of values. This point is often overlooked by beginning students, who simply assert their uncertainties but omit any justification. Without a brief explanation of how the uncertainty was estimated, the assertion is almost useless.

The most important point about our two experts' measurements is this: Like most scientific measurements, they would both have been useless if they had not included reliable statements of their uncertainties. In fact, if we knew only the two best estimates (15 for George and 13.9 for Martha), not only would we have been unable to draw a valid conclusion, but we could actually have been misled, because George's result (15) seems to suggest the crown is genuine.

1.4 More Examples

The examples in the past two sections were chosen, not for their great importance, but to introduce some principal features of error analysis. Thus, you can be excused for thinking them a little contrived. It is easy, however, to think of examples of great importance in almost any branch of applied or basic science.

In the applied sciences, for example, the engineers designing a power plant must know the characteristics of the materials and fuels they plan to use. The manufacturer of a pocket calculator must know the properties of its various electronic components. In each case, somebody must measure the required parameters, and having measured them, must establish their reliability, which requires error analysis. Engineers concerned with the safety of airplanes, trains, or cars must understand the uncertainties in drivers' reaction times, in braking distances, and in a host of other variables; failure to carry out error analysis can lead to accidents such as that shown on the cover of this book. Even in a less scientific field, such as the manufacture of clothing, error analysis in the form of quality control plays a vital part.

In the basic sciences, error analysis has an even more fundamental role. When any new theory is proposed, it must be tested against older theories by means of one or more experiments for which the new and old theories predict different outcomes. In principle, a researcher simply performs the experiment and lets the outcome decide between the rival theories. In practice, however, the situation is complicated by the inevitable experimental uncertainties. These uncertainties must all be analyzed carefully and their effects reduced until the experiment singles out one acceptable theory. That is, the experimental results, with their uncertainties, must be *consistent* with the predictions of one theory and *inconsistent* with those of all known, reasonable alternatives. Obviously, the success of such a procedure depends critically on the scientist's understanding of error analysis and ability to convince others of this understanding.

A famous example of such a test of a scientific theory is the measurement of the bending of light as it passes near the sun. When Einstein published his general theory of relativity in 1916, he pointed out that the theory predicted that light from a star would be bent through an angle $\alpha = 1.8''$ as it passes near the sun. The simplest classical theory would predict no bending ($\alpha = 0$), and a more careful classical analysis would predict (as Einstein himself noted in 1911) bending through an angle $\alpha = 0.9''$. In principle, all that was necessary was to observe a star when it was aligned with the edge of the sun and to measure the angle of bending α . If the result were $\alpha = 1.8''$, general relativity would be vindicated (at least for this phenomenon); if α were found to be 0 or $0.9''$, general relativity would be wrong and one of the older theories right.

In practice, measuring the bending of light by the sun was extremely hard and was possible only during a solar eclipse. Nonetheless, in 1919 it was successfully measured by Dyson, Eddington, and Davidson, who reported their best estimate as $\alpha = 2''$, with 95% confidence that it lay between $1.7''$ and $2.3''$.¹ Obviously, this result was consistent with general relativity and inconsistent with either of the older predictions. Therefore, it gave strong support to Einstein's theory of general relativity.

At the time, this result was controversial. Many people suggested that the uncertainties had been badly underestimated and hence that the experiment was inconclusive. Subsequent experiments have tended to confirm Einstein's prediction and to vindicate the conclusion of Dyson, Eddington, and Davidson. The important point here is that the whole question hinged on the experimenters' ability to estimate reliably all their uncertainties and to convince everyone else they had done so.

Students in introductory physics laboratories are not usually able to conduct definitive tests of new theories. Often, however, they do perform experiments that test existing physical theories. For example, Newton's theory of gravity predicts that bodies fall with constant acceleration g (under the appropriate conditions), and students can conduct experiments to test whether this prediction is correct. At first, this kind of experiment may seem artificial and pointless because the theories have obvi-

¹This simplified account is based on the original paper of F. W. Dyson, A. S. Eddington, and C. Davidson (*Philosophical Transactions of the Royal Society*, **220A**, 1920, 291). I have converted the *probable error* originally quoted into the 95% confidence limits. The precise significance of such confidence limits will be established in Chapter 5.

ously been tested many times with much more precision than possible in a teaching laboratory. Nonetheless, if you understand the crucial role of error analysis and accept the challenge to make the most precise test possible with the available equipment, such experiments can be interesting and instructive exercises.

1.5 Estimating Uncertainties When Reading Scales

Thus far, we have considered several examples that illustrate why every measurement suffers from uncertainties and why their magnitude is important to know. We have not yet discussed how we can actually evaluate the magnitude of an uncertainty. Such evaluation can be fairly complicated and is the main topic of this book. Fortunately, reasonable estimates of the uncertainty of some simple measurements are easy to make, often using no more than common sense. Here and in Section 1.6, I discuss examples of such measurements. An understanding of these examples will allow you to begin using error analysis in your experiments and will form the basis for later discussions.

The first example is a measurement using a marked scale, such as the ruler in Figure 1.2 or the voltmeter in Figure 1.3. To measure the length of the pencil in

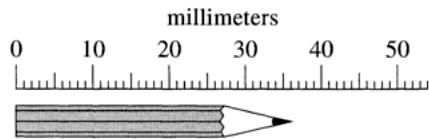


Figure 1.2. Measuring a length with a ruler.

Figure 1.2, we must first place the end of the pencil opposite the zero of the ruler and then decide where the tip comes to on the ruler's scale. To measure the voltage in Figure 1.3, we have to decide where the needle points on the voltmeter's scale. If we assume the ruler and voltmeter are reliable, then in each case the main prob-

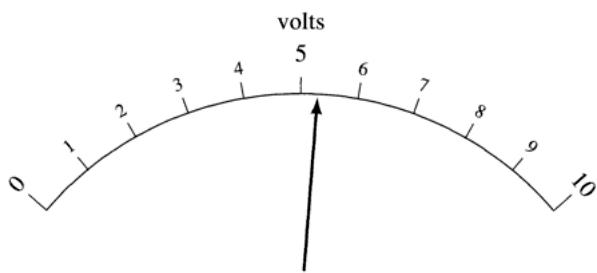


Figure 1.3. A reading on a voltmeter.

lem is to decide where a certain point lies in relation to the scale markings. (Of course, if there is any possibility the ruler and voltmeter are *not* reliable, we will have to take this uncertainty into account as well.)

The markings of the ruler in Figure 1.2 are fairly close together (1 mm apart). We might reasonably decide that the length shown is undoubtedly closer to 36 mm than it is to 35 or 37 mm but that no more precise reading is possible. In this case, we would state our conclusion as

$$\begin{aligned} \text{best estimate of length} &= 36 \text{ mm}, \\ \text{probable range: } &35.5 \text{ to } 36.5 \text{ mm} \end{aligned} \quad (1.1)$$

and would say that we have measured the length to the nearest millimeter.

This type of conclusion—that the quantity lies closer to a given mark than to either of its neighboring marks—is quite common. For this reason, many scientists introduce the convention that the statement “ $l = 36 \text{ mm}$ ” without any qualification is presumed to mean that l is closer to 36 than to 35 or 37; that is,

$$l = 36 \text{ mm}$$

means

$$35.5 \text{ mm} \leq l \leq 36.5 \text{ mm}.$$

In the same way, an answer such as $x = 1.27$ without any stated uncertainty would be presumed to mean that x lies between 1.265 and 1.275. In this book, I do not use this convention but instead always indicate uncertainties explicitly. Nevertheless, you need to understand the convention and know that it applies to any number stated without an uncertainty, especially in this age of pocket calculators, which display many digits. If you unthinkingly copy a number such as 123.456 from your calculator without any qualification, then your reader is entitled to assume the number is definitely correct to six significant figures, which is very unlikely.

The markings on the voltmeter shown in Figure 1.3 are more widely spaced than those on the ruler. Here, most observers would agree that you can do better than simply identify the mark to which the pointer is closest. Because the spacing is larger, you can realistically estimate where the pointer lies in the space between two marks. Thus, a reasonable conclusion for the voltage shown might be

$$\begin{aligned} \text{best estimate of voltage} &= 5.3 \text{ volts}, \\ \text{probable range: } &5.2 \text{ to } 5.4 \text{ volts.} \end{aligned} \quad (1.2)$$

The process of estimating positions between the scale markings is called *interpolation*. It is an important technique that can be improved with practice.

Different observers might not agree with the precise estimates given in Equations (1.1) and (1.2). You might well decide that you could interpolate for the length in Figure 1.2 and measure it with a smaller uncertainty than that given in Equation (1.1). Nevertheless, few people would deny that Equations (1.1) and (1.2) are *reasonable* estimates of the quantities concerned and of their probable uncertainties. Thus, we see that approximate estimation of uncertainties is fairly easy when the only problem is to locate a point on a marked scale.

1.6 Estimating Uncertainties in Repeatable Measurements

Many measurements involve uncertainties that are much harder to estimate than those connected with locating points on a scale. For example, when we measure a time interval using a stopwatch, the main source of uncertainty is not the difficulty of reading the dial but our own unknown reaction time in starting and stopping the watch. Sometimes these kinds of uncertainty can be estimated reliably, if we can repeat the measurement several times. Suppose, for example, we time the period of a pendulum once and get an answer of 2.3 seconds. From one measurement, we can't say much about the experimental uncertainty. But if we repeat the measurement and get 2.4 seconds, then we can immediately say that the uncertainty is probably of the order of 0.1 s. If a sequence of four timings gives the results (in seconds),

$$2.3, 2.4, 2.5, 2.4, \quad (1.3)$$

then we can begin to make some fairly realistic estimates.

First, a natural assumption is that the best estimate of the period is the *average² value*, 2.4 s.

Second, another reasonably safe assumption is that the correct period lies between the lowest value, 2.3, and the highest, 2.5. Thus, we might reasonably conclude that

$$\begin{aligned} \text{best estimate} &= \text{average} = 2.4 \text{ s}, \\ \text{probable range: } &2.3 \text{ to } 2.5 \text{ s.} \end{aligned} \quad (1.4)$$

Whenever you can repeat the same measurement several times, the spread in your measured values gives a valuable indication of the uncertainty in your measurements. In Chapters 4 and 5, I discuss statistical methods for treating such repeated measurements. Under the right conditions, these statistical methods give a more accurate estimate of uncertainty than we have found in Equation (1.4) using just common sense. A proper statistical treatment also has the advantage of giving an objective value for the uncertainty, independent of the observer's individual judgment.³ Nevertheless, the estimate in statement (1.4) represents a simple, realistic conclusion to draw from the four measurements in (1.3).

Repeated measurements such as those in (1.3) cannot always be relied on to reveal the uncertainties. First, we must be sure that the quantity measured is really the *same* quantity each time. Suppose, for example, we measure the breaking strength of two supposedly identical wires by breaking them (something we can't do more than once with each wire). If we get two different answers, this difference *may* indicate that our measurements were uncertain *or* that the two wires were not really identical. By itself, the difference between the two answers sheds no light on the reliability of our measurements.

²I will prove in Chapter 5 that the best estimate based on several measurements of a quantity is almost always the average of the measurements.

³Also, a proper statistical treatment usually gives a *smaller* uncertainty than the full range from the lowest to the highest observed value. Thus, upon looking at the four timings in (1.3), we have judged that the period is "probably" somewhere between 2.3 and 2.5 s. The statistical methods of Chapters 4 and 5 let us state with 70% confidence that the period lies in the smaller range of 2.36 to 2.44 s.

Even when we can be sure we are measuring the same quantity each time, repeated measurements do not always reveal uncertainties. For example, suppose the clock used for the timings in (1.3) was running consistently 5% fast. Then, all timings made with it will be 5% too long, and no amount of repeating (with the same clock) will reveal this deficiency. Errors of this sort, which affect all measurements in the same way, are called *systematic* errors and can be hard to detect, as discussed in Chapter 4. In this example, the remedy is to check the clock against a more reliable one. More generally, if the reliability of any measuring device is in doubt, it should clearly be checked against a device known to be more reliable.

The examples discussed in this and the previous section show that experimental uncertainties sometimes can be estimated easily. On the other hand, many measurements have uncertainties that are *not* so easily evaluated. Also, we ultimately want more precise values for the uncertainties than the simple estimates just discussed. These topics will occupy us from Chapter 3 onward. In Chapter 2, I assume temporarily that you know how to estimate the uncertainties in all quantities of interest, so that we can discuss how the uncertainties are best reported and how they are used in drawing an experimental conclusion.

Chapter 2

How to Report and Use Uncertainties

Having read Chapter 1, you should now have some idea of the importance of experimental uncertainties and how they arise. You should also understand how uncertainties can be estimated in a few simple situations. In this chapter, you will learn some basic notations and rules of error analysis and study examples of their use in typical experiments in a physics laboratory. The aim is to familiarize you with the basic vocabulary of error analysis and its use in the introductory laboratory. Chapter 3 begins a systematic study of how uncertainties are actually evaluated.

Sections 2.1 to 2.3 define several basic concepts in error analysis and discuss general rules for stating uncertainties. Sections 2.4 to 2.6 discuss how these ideas could be used in typical experiments in an introductory physics laboratory. Finally, Sections 2.7 to 2.9 introduce fractional uncertainty and discuss its significance.

2.1 Best Estimate \pm Uncertainty

We have seen that the correct way to state the result of measurement is to give a best estimate of the quantity and the range within which you are confident the quantity lies. For example, the result of the timings discussed in Section 1.6 was reported as

$$\begin{aligned} \text{best estimate of time} &= 2.4 \text{ s}, \\ \text{probable range: } &2.3 \text{ to } 2.5 \text{ s.} \end{aligned} \tag{2.1}$$

Here, the best estimate, 2.4 s, lies at the midpoint of the estimated range of probable values, 2.3 to 2.5 s, as it has in all the examples. This relationship is obviously natural and pertains in most measurements. It allows the results of the measurement to be expressed in compact form. For example, the measurement of the time recorded in (2.1) is usually stated as follows:

$$\text{measured value of time} = 2.4 \pm 0.1 \text{ s.} \tag{2.2}$$

This single equation is equivalent to the two statements in (2.1).

In general, the result of any measurement of a quantity x is stated as

$$(\text{measured value of } x) = x_{\text{best}} \pm \delta x. \tag{2.3}$$

This statement means, first, that the experimenter's best estimate for the quantity concerned is the number x_{best} , and second, that he or she is reasonably confident the quantity lies somewhere between $x_{\text{best}} - \delta x$ and $x_{\text{best}} + \delta x$. The number δx is called the *uncertainty*, or *error*, or *margin of error* in the measurement of x . For convenience, the uncertainty δx is always defined to be positive, so that $x_{\text{best}} + \delta x$ is always the *highest* probable value of the measured quantity and $x_{\text{best}} - \delta x$ the *lowest*.

I have intentionally left the meaning of the range $x_{\text{best}} - \delta x$ to $x_{\text{best}} + \delta x$ somewhat vague, but it can sometimes be made more precise. In a simple measurement such as that of the height of a doorway, we can easily state a range $x_{\text{best}} - \delta x$ to $x_{\text{best}} + \delta x$ within which we are *absolutely* certain the measured quantity lies. Unfortunately, in most scientific measurements, such a statement is hard to make. In particular, to be *completely* certain that the measured quantity lies between $x_{\text{best}} - \delta x$ and $x_{\text{best}} + \delta x$, we usually have to choose a value for δx that is too large to be useful. To avoid this situation, we can sometimes choose a value for δx that lets us state with a certain percent confidence that the actual quantity lies within the range $x_{\text{best}} \pm \delta x$. For instance, the public opinion polls conducted during elections are traditionally stated with margins of error that represent 95% confidence limits. The statement that 60% of the electorate favor Candidate A, with a margin of error of 3 percentage points (60 ± 3), means that the pollsters are 95% confident that the percent of voters favoring Candidate A is between 57 and 63; in other words, after many elections, we should expect the correct answer to have been *inside* the stated margins of error 95% of the times and *outside* these margins only 5% of the times.

Obviously, we cannot state a percent confidence in our margins of error until we understand the statistical laws that govern the process of measurement. I return to this point in Chapter 4. For now, let us be content with defining the uncertainty δx so that we are "reasonably certain" the measured quantity lies between $x_{\text{best}} - \delta x$ and $x_{\text{best}} + \delta x$.

- Quick Check¹ 2.1.** (a) A student measures the length of a simple pendulum and reports his best estimate as 110 mm and the range in which the length probably lies as 108 to 112 mm. Rewrite this result in the standard form (2.3). (b) If another student reports her measurement of a current as $I = 3.05 \pm 0.03$ amps, what is the range within which I probably lies?

2.2 Significant Figures

Several basic rules for stating uncertainties are worth emphasizing. First, because the quantity δx is an estimate of an uncertainty, obviously it should not be stated

¹These "Quick Checks" appear at intervals through the text to give you a chance to check your understanding of the concept just introduced. They are straightforward exercises, and many can be done in your head. I urge you to take a moment to make sure you can do them; if you cannot, you should reread the preceding few paragraphs.

with too much precision. If we measure the acceleration of gravity g , it would be absurd to state a result like

$$(\text{measured } g) = 9.82 \pm 0.02385 \text{ m/s}^2. \quad (2.4)$$

The uncertainty in the measurement cannot conceivably be known to four significant figures. In high-precision work, uncertainties are sometimes stated with two significant figures, but for our purposes we can state the following rule:

Rule for Stating Uncertainties

Experimental uncertainties should almost always be rounded to one significant figure. (2.5)

Thus, if some calculation yields the uncertainty $\delta g = 0.02385 \text{ m/s}^2$, this answer should be rounded to $\delta g = 0.02 \text{ m/s}^2$, and the conclusion (2.4) should be rewritten as

$$(\text{measured } g) = 9.82 \pm 0.02 \text{ m/s}^2. \quad (2.6)$$

An important practical consequence of this rule is that many error calculations can be carried out mentally without using a calculator or even pencil and paper.

The rule (2.5) has only one significant exception. If the leading digit in the uncertainty δx is a 1, then keeping two significant figures in δx may be better. For example, suppose that some calculation gave the uncertainty $\delta x = 0.14$. Rounding this number to $\delta x = 0.1$ would be a substantial proportionate reduction, so we could argue that retaining two figures might be less misleading, and quote $\delta x = 0.14$. The same argument could perhaps be applied if the leading digit is a 2 but certainly not if it is any larger.

Once the uncertainty in a measurement has been estimated, the significant figures in the measured value must be considered. A statement such as

$$\text{measured speed} = 6051.78 \pm 30 \text{ m/s} \quad (2.7)$$

is obviously ridiculous. The uncertainty of 30 means that the digit 5 might really be as small as 2 or as large as 8. Clearly the trailing digits 1, 7, and 8 have no significance at all and should be rounded. That is, the correct statement of (2.7) is

$$\text{measured speed} = 6050 \pm 30 \text{ m/s}. \quad (2.8)$$

The general rule is this:

Rule for Stating Answers

The last significant figure in any stated answer should usually be of the same order of magnitude (in the same decimal position) as the uncertainty. (2.9)

For example, the answer 92.81 with an uncertainty of 0.3 should be rounded as

$$92.8 \pm 0.3.$$

If its uncertainty is 3, then the same answer should be rounded as

$$93 \pm 3,$$

and if the uncertainty is 30, then the answer should be

$$90 \pm 30.$$

An important qualification to rules (2.5) and (2.9) is as follows: To reduce inaccuracies caused by rounding, *any numbers to be used in subsequent calculations should normally retain at least one significant figure more than is finally justified*. At the end of the calculations, the final answer should be rounded to remove these extra, insignificant figures. An electronic calculator will happily carry numbers with far more digits than are likely to be significant in any calculation you make in a laboratory. Obviously, these numbers do not need to be rounded in the middle of a calculation but certainly must be rounded appropriately for the final answers.²

Note that the uncertainty in any measured quantity has the same dimensions as the measured quantity itself. Therefore, writing the units (m/s^2 , cm^3 , etc.) after both the answer *and* the uncertainty is clearer and more economical, as in Equations (2.6) and (2.8). By the same token, if a measured number is so large or small that it calls for scientific notation (the use of the form 3×10^3 instead of 3,000, for example), then it is simpler and clearer to put the answer and uncertainty in the same form. For example, the result

$$\text{measured charge} = (1.61 \pm 0.05) \times 10^{-19} \text{ coulombs}$$

is much easier to read and understand in this form than it would be in the form

$$\text{measured charge} = 1.61 \times 10^{-19} \pm 5 \times 10^{-21} \text{ coulombs.}$$

Quick Check 2.2. Rewrite each of the following measurements in its most appropriate form:

- (a) $v = 8.123456 \pm 0.0312 \text{ m/s}$
- (b) $x = 3.1234 \times 10^4 \pm 2 \text{ m}$
- (c) $m = 5.6789 \times 10^{-7} \pm 3 \times 10^{-9} \text{ kg.}$

2.3 Discrepancy

Before I address the question of how to use uncertainties in experimental reports, a few important terms should be introduced and defined. First, if two measurements

²Rule (2.9) has one more small exception. If the leading digit in the uncertainty is small (a 1 or, perhaps, a 2), retaining one extra digit in the final answer may be appropriate. For example, an answer such as 3.6 ± 1 is quite acceptable because one could argue that rounding it to 4 ± 1 would waste information.

of the same quantity disagree, we say there is a *discrepancy*. Numerically, we define the discrepancy between two measurements as their difference:

$$\text{discrepancy} = \text{difference between two measured values of the same quantity.} \quad (2.10)$$

More specifically, each of the two measurements consists of a best estimate and an uncertainty, and we define the discrepancy as the difference between the two best estimates. For example, if two students measure the same resistance as follows

Student A: 15 ± 1 ohms

and

Student B: 25 ± 2 ohms,

their discrepancy is

$$\text{discrepancy} = 25 - 15 = 10 \text{ ohms.}$$

Recognize that a discrepancy may or may not be *significant*. The two measurements just discussed are illustrated in Figure 2.1(a), which shows clearly that the discrepancy of 10 ohms is *significant* because no single value of the resistance is compatible with both measurements. Obviously, at least one measurement is incorrect, and some careful checking is needed to find out what went wrong.

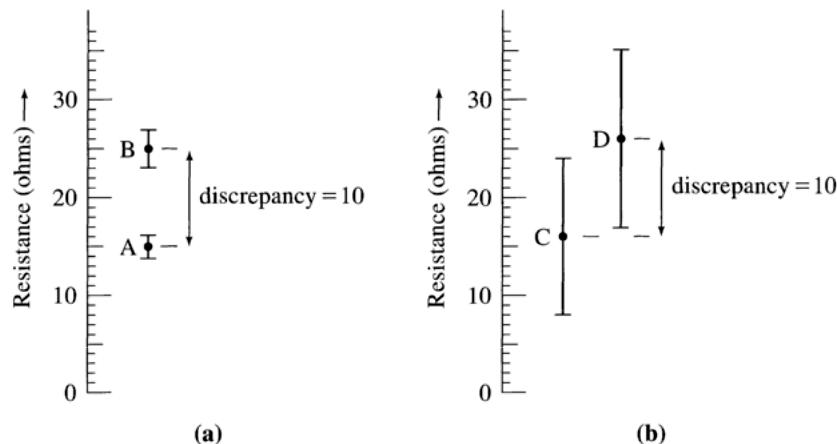


Figure 2.1. (a) Two measurements of the same resistance. Each measurement includes a best estimate, shown by a block dot, and a range of probable values, shown by a vertical error bar. The discrepancy (difference between the two best estimates) is 10 ohms and is *significant* because it is much larger than the combined uncertainty in the two measurements. Almost certainly, at least one of the experimenters made a mistake. (b) Two different measurements of the same resistance. The discrepancy is again 10 ohms, but in this case it is *insignificant* because the stated margins of error overlap. There is no reason to doubt either measurement (although they could be criticized for being rather imprecise).

Suppose, on the other hand, two other students had reported these results:

Student C: 16 ± 8 ohms

and

Student D: 26 ± 9 ohms.

Here again, the discrepancy is 10 ohms, but in this case the discrepancy is *insignificant* because, as shown in Figure 2.1(b), the two students' margins of error overlap comfortably and both measurements could well be correct. The discrepancy between two measurements of the same quantity should be assessed not just by its size but, more importantly, by how big it is *compared with the uncertainties in the measurements*.

In the teaching laboratory, you may be asked to measure a quantity that has been measured carefully many times before, and for which an accurate *accepted value* is known and published, for example, the electron's charge or the universal gas constant. This accepted value is not exact, of course; it is the result of measurements and, like all measurements, has some uncertainty. Nonetheless, in many cases the accepted value is much more accurate than you could possibly achieve yourself. For example, the currently accepted value of the universal gas constant R is

$$(\text{accepted } R) = 8.31451 \pm 0.00007 \text{ J/(mol}\cdot\text{K).} \quad (2.11)$$

As expected, this value *is* uncertain, but the uncertainty is extremely small by the standards of most teaching laboratories. Thus, when you compare your measured value of such a constant with the accepted value, you can usually treat the accepted value as exact.³

Although many experiments call for measurement of a quantity whose accepted value is known, few require measurement of a quantity whose *true* value is known.⁴ In fact, the true value of a measured quantity can almost *never* be known exactly and is, in fact, hard to define. Nevertheless, discussing the difference between a measured value and the corresponding true value is sometimes useful. Some authors call this difference the *true error*.

2.4 Comparison of Measured and Accepted Values

Performing an experiment without drawing some sort of conclusion has little merit. A few experiments may have mainly qualitative results—the appearance of an interference pattern on a ripple tank or the color of light transmitted by some optical system—but the vast majority of experiments lead to *quantitative* conclusions, that is, to a statement of numerical results. It is important to recognize that the statement of a *single measured number* is completely uninteresting. Statements that the density

³This is not always so. For example, if you look up the refractive index of glass, you find values ranging from 1.5 to 1.9, depending on the composition of the glass. In an experiment to measure the refractive index of a piece of glass whose composition is unknown, the accepted value is therefore no more than a rough guide to the expected answer.

⁴Here is an example: If you measure the ratio of a circle's circumference to its diameter, the true answer is exactly π . (Obviously such an experiment is rather contrived.)

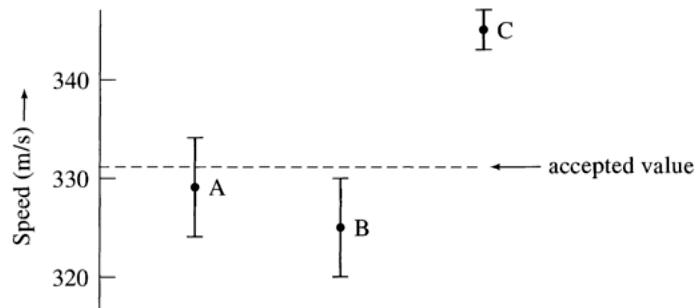


Figure 2.2. Three measurements of the speed of sound at standard temperature and pressure. Because the accepted value (331 m/s) is within Student A's margins of error, her result is satisfactory. The accepted value is just outside Student B's margin of error, but his measurement is nevertheless acceptable. The accepted value is *far outside* Student C's stated margins, and his measurement is definitely unsatisfactory.

of some metal was measured as 9.3 ± 0.2 gram/cm³ or that the momentum of a cart was measured as 0.051 ± 0.004 kg·m/s are, by themselves, of no interest. An interesting conclusion must *compare two or more numbers*: a measurement with the accepted value, a measurement with a theoretically predicted value, or several measurements, to show that they are related to one another in accordance with some physical law. It is in such comparison of numbers that error analysis is so important. This and the next two sections discuss three typical experiments to illustrate how the estimated uncertainties are used to draw a conclusion.

Perhaps the simplest type of experiment is a measurement of a quantity whose accepted value is known. As discussed, this exercise is a somewhat artificial experiment peculiar to the teaching laboratory. The procedure is to measure the quantity, estimate the experimental uncertainty, and compare these values with the accepted value. Thus, in an experiment to measure the speed of sound in air (at standard temperature and pressure), Student A might arrive at the conclusion

$$\text{A's measured speed} = 329 \pm 5 \text{ m/s}, \quad (2.12)$$

compared with the

$$\text{accepted speed} = 331 \text{ m/s}. \quad (2.13)$$

Student A might choose to display this result graphically as in Figure 2.2. She should certainly include in her report both Equations (2.12) and (2.13) next to each other, so her readers can clearly appreciate her result. She should probably add an explicit statement that because the accepted value lies inside her margins of error, her measurement seems satisfactory.

The meaning of the uncertainty δx is that the correct value of x *probably* lies between $x_{\text{best}} - \delta x$ and $x_{\text{best}} + \delta x$; it is certainly *possible* that the correct value lies slightly outside this range. Therefore, a measurement can be regarded as satisfactory even if the accepted value lies slightly outside the estimated range of the measured

value. For example, if Student B found the value

$$\text{B's measured speed} = 325 \pm 5 \text{ m/s},$$

he could certainly claim that his measurement is consistent with the accepted value of 331 m/s.

On the other hand, if the accepted value is well outside the margins of error (the discrepancy is appreciably more than twice the uncertainty, say), there is reason to think something has gone wrong. For example, suppose the unlucky Student C finds

$$\text{C's measured speed} = 345 \pm 2 \text{ m/s} \quad (2.14)$$

compared with the

$$\text{accepted speed} = 331 \text{ m/s}. \quad (2.15)$$

Student C's discrepancy is 14 m/s, which is seven times bigger than his stated uncertainty (see Figure 2.2). He will need to check his measurements and calculations to find out what has gone wrong.

Unfortunately, the tracing of C's mistake may be a tedious business because of the numerous possibilities. He may have made a mistake in the measurements or calculations that led to the answer 345 m/s. He may have estimated his uncertainty incorrectly. (The answer 345 ± 15 m/s would have been acceptable.) He also might be comparing his measurement with the wrong accepted value. For example, the accepted value 331 m/s is the speed of sound at standard temperature and pressure. Because standard temperature is 0°C, there is a good chance the measured speed in (2.14) was *not* taken at standard temperature. In fact, if the measurement was made at 20°C (that is, normal room temperature), the correct accepted value for the speed of sound is 343 m/s, and the measurement would be entirely acceptable.

Finally, and perhaps most likely, a discrepancy such as that between (2.14) and (2.15) may indicate some undetected source of systematic error (such as a clock that runs consistently slow, as discussed in Chapter 1). Detection of such systematic errors (ones that consistently push the result in one direction) requires careful checking of the calibration of all instruments and detailed review of all procedures.

2.5 Comparison of Two Measured Numbers

Many experiments involve measuring two numbers that theory predicts should be equal. For example, the law of conservation of momentum states that the total momentum of an isolated system is constant. To test it, we might perform a series of experiments with two carts that collide as they move along a frictionless track. We could measure the total momentum of the two carts before (p) and after (q) they collide and check whether $p=q$ within experimental uncertainties. For a single pair of measurements, our results could be

$$\text{initial momentum } p = 1.49 \pm 0.03 \text{ kg}\cdot\text{m/s}$$

and

$$\text{final momentum } q = 1.56 \pm 0.06 \text{ kg}\cdot\text{m/s}.$$

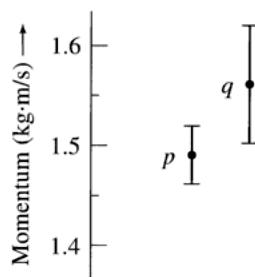


Figure 2.3. Measured values of the total momentum of two carts before (p) and after (q) a collision. Because the margins of error for p and q overlap, these measurements are certainly consistent with conservation of momentum (which implies that p and q should be equal).

Here, the range in which p probably lies (1.46 to 1.52) *overlaps* the range in which q probably lies (1.50 to 1.62). (See Figure 2.3.) Therefore, these measurements are consistent with conservation of momentum. If, on the other hand, the two probable ranges were not even close to overlapping, the measurements would be inconsistent with conservation of momentum, and we would have to check for mistakes in our measurements or calculations, for possible systematic errors, and for the possibility that some external forces (such as gravity or friction) are causing the momentum of the system to change.

If we repeat similar pairs of measurements several times, what is the best way to display our results? First, using a table to record a sequence of similar measurements is usually better than listing the results as several distinct statements. Second, the uncertainties often differ little from one measurement to the next. For example, we might convince ourselves that the uncertainties in all measurements of the initial momentum p are about $\delta p \approx 0.03$ kg·m/s and that the uncertainties in the final q are all about $\delta q \approx 0.06$ kg·m/s. If so, a good way to display our measurements would be as shown in Table 2.1.

Table 2.1. Measured momenta (kg·m/s).

Trial number	Initial momentum p (all ± 0.03)	Final momentum q (all ± 0.06)
1	1.49	1.56
2	3.10	3.12
3	2.16	2.05
etc.		

For each pair of measurements, the probable range of values for p overlaps (or nearly overlaps) the range of values for q . If this overlap continues for all measurements, our results can be pronounced consistent with conservation of momentum. Note that our experiment does not *prove* conservation of momentum; no experiment can. The best you can hope for is to conduct many more trials with progressively

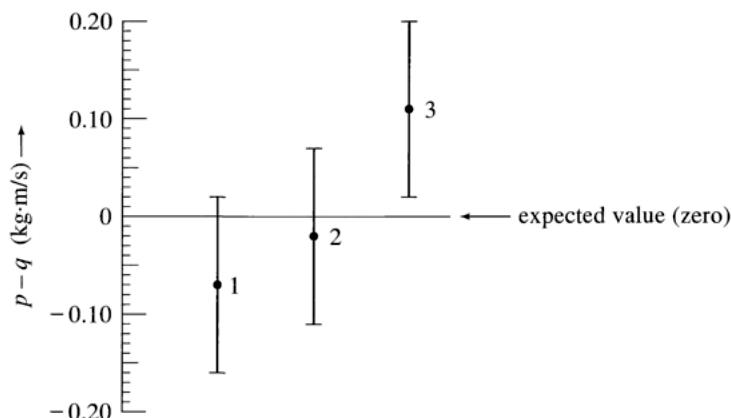


Figure 2.4. Three trials in a test of the conservation of momentum. The student has measured the total momentum of two carts before and after they collide (p and q , respectively). If momentum is conserved, the differences $p - q$ should all be zero. The plot shows the value of $p - q$ with its error bar for each trial. The expected value 0 is inside the margins of error in trials 1 and 2 and only slightly outside in trial 3. Therefore, these results are consistent with the conservation of momentum.

Whether our results are consistent with conservation of momentum can now be seen at a glance by checking whether the numbers in the final column are consistent with zero (that is, are less than, or comparable with, the uncertainty 0.09). Alternatively, and perhaps even better, we could plot the results as in Figure 2.4 and check visually. Yet another way to achieve the same effect would be to calculate the *ratios* q/p , which should all be consistent with the expected value $q/p = 1$. (Here, we would need to calculate the uncertainty in q/p , a problem discussed in Chapter 3.)

Our discussion of the uncertainty in $p - q$ applies to the difference of any two measured numbers. If we had measured any two numbers x and y and used our measured values to compute the difference $x - y$, by the argument just given, the resulting uncertainty in the difference would be the *sum* of the separate uncertainties in x and y . We have, therefore, established the following provisional rule:

**Uncertainty in a Difference
(Provisional Rule)**

If two quantities x and y are measured with uncertainties δx and δy , and if the measured values x and y are used to calculate the difference $q = x - y$, the *uncertainty in q* is the *sum of the uncertainties in x and y* :

$$\delta q \approx \delta x + \delta y. \quad (2.18)$$

I call this rule “provisional” because we will find in Chapter 3 that the uncertainty in the quantity $q = x - y$ is often somewhat smaller than that given by Equation

(2.18). Thus, we will be replacing the provisional rule (2.18) by an “improved” rule—in which the uncertainty in $q = x - y$ is given by the so-called quadratic sum of δx and δy , as defined in Equation (3.13). Because this improved rule gives a somewhat smaller uncertainty for q , you will want to use it when appropriate. For now, however, let us be content with the provisional rule (2.18) for three reasons: (1) The rule (2.18) is easy to understand—much more so than the improved rule of Chapter 3. (2) In most cases, the difference between the two rules is small. (3) The rule (2.18) always gives an upper bound on the uncertainty in $q = x - y$; thus, we know at least that the uncertainty in $x - y$ is never worse than the answer given in (2.18).

The result (2.18) is the first in a series of rules for the *propagation of errors*. To calculate a quantity q in terms of measured quantities x and y , we need to know how the uncertainties in x and y “propagate” to cause uncertainty in q . A complete discussion of error propagation appears in Chapter 3.

Quick Check 2.3. In an experiment to measure the latent heat of ice, a student adds a chunk of ice to water in a styrofoam cup and observes the change in temperature as the ice melts. To determine the mass of ice added, she weighs the cup of water before and after she adds the ice and then takes the difference. If her two measurements were

$$(\text{mass of cup \& water}) = m_1 = 203 \pm 2 \text{ grams}$$

and

$$(\text{mass of cup, water, \& ice}) = m_2 = 246 \pm 3 \text{ grams},$$

find her answer for the mass of ice, $m_2 - m_1$, with its uncertainty, as given by the provisional rule (2.18).

2.6 Checking Relationships with a Graph

Many physical laws imply that one quantity should be proportional to another. For example, Hooke’s law states that the extension of a spring is proportional to the force stretching it, and Newton’s law says that the acceleration of a body is proportional to the total applied force. Many experiments in a teaching laboratory are designed to check this kind of proportionality.

If one quantity y is proportional to some other quantity x , a graph of y against x is a straight line through the origin. Thus, to test whether y is proportional to x , you can plot the measured values of y against those of x and note whether the resulting points do lie on a straight line through the origin. Because a straight line is so easily recognizable, this method is a simple, effective way to check for proportionality.

To illustrate this use of graphs, let us imagine an experiment to test Hooke’s law. This law, usually written as $F = kx$, asserts that the extension x of a spring is proportional to the force F stretching it, so $x = F/k$, where k is the “force constant”

of the spring. A simple way to test this law is to hang the spring vertically and suspend various masses m from it. Here, the force F is the weight mg of the load; so the extension should be

$$x = \frac{mg}{k} = \left(\frac{g}{k}\right)m. \quad (2.19)$$

The extension x should be proportional to the load m , and a graph of x against m should be a straight line through the origin.

If we measure x for a variety of different loads m and plot our measured values of x and m , the resulting points almost certainly will not lie *exactly* on a straight line. Suppose, for example, we measure the extension x for eight different loads m and get the results shown in Table 2.3. These values are plotted in Figure 2.5(a),

Table 2.3. Load and extension.

Load m (grams) (δm negligible)	200	300	400	500	600	700	800	900
Extension x (cm) (all ± 0.3)	1.1	1.5	1.9	2.8	3.4	3.5	4.6	5.4

which also shows a possible straight line that passes through the origin and is reasonably close to all eight points. As we should have expected, the eight points do not lie exactly on any line. The question is whether this result stems from experimental uncertainties (as we would hope), from mistakes we have made, or even from the possibility the extension x is *not* proportional to m . To answer this question, we must consider our uncertainties.

As usual, the measured quantities, extensions x and masses m , are subject to uncertainty. For simplicity, let us suppose that the masses used are known very accurately, so that the uncertainty in m is negligible. Suppose, on the other hand, that all measurements of x have an uncertainty of approximately 0.3 cm (as indicated in Table 2.3). For a load of 200 grams, for example, the extension would probably be in the range 1.1 ± 0.3 cm. Our first experimental point on the graph thus lies on the vertical line $m = 200$ grams, somewhere between $x = 0.8$ and $x = 1.4$ cm. This range is indicated in Figure 2.5(b), which shows an *error bar* through each point to indicate the range in which it probably lies. Obviously, we should expect to find a straight line that goes through the origin and *passes through or close to all the error bars*. Figure 2.5(b) has such a line, so we conclude that the data on which Figure 2.5(b) is based are consistent with x being proportional to m .

We saw in Equation (2.19) that the slope of the graph of x against m is g/k . By measuring the slope of the line in Figure 2.5(b), we can therefore find the constant k of the spring. By drawing the steepest and least steep lines that fit the data reasonably well, we could also find the uncertainty in this value for k . (See Problem 2.18.)

If the best straight line misses a high proportion of the error bars or if it misses any by a large distance (compared with the length of the error bars), our results

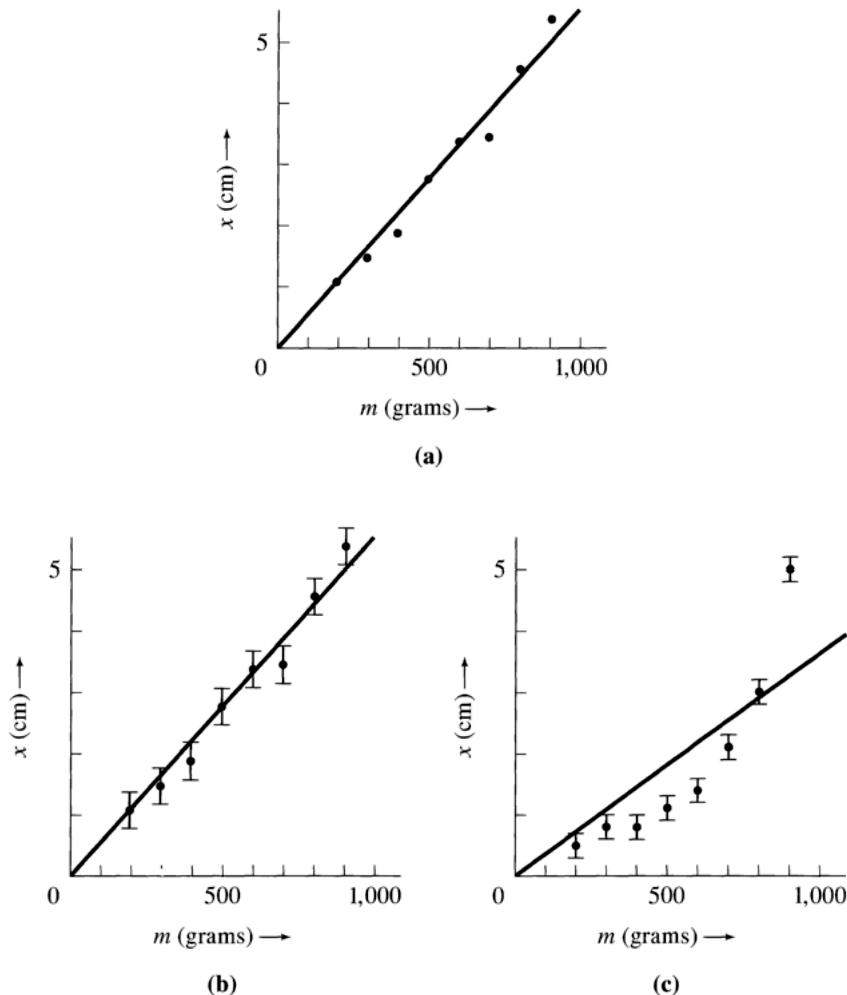


Figure 2.5. Three plots of extension x of a spring against the load m . (a) The data of Table 2.3 without error bars. (b) The same data with error bars to show the uncertainties in x . (The uncertainties in m are assumed to be negligible.) These data are consistent with the expected proportionality of x and m . (c) A different set of data, which are inconsistent with x being proportional to m .

would be *inconsistent* with x being proportional to m . This situation is illustrated in Figure 2.5(c). With the results shown there, we would have to recheck our measurements and calculations (including the calculation of the uncertainties) and consider whether x is *not* proportional to m for some reason. [In Figure 2.5(c), for instance, the first five points can be fitted to a straight line through the origin. This situation suggests that x may be proportional to m up to approximately 600 grams, but that Hooke's law breaks down at that point and the spring starts to stretch more rapidly.]

Thus far, we have supposed that the uncertainty in the mass (which is plotted along the horizontal axis) is negligible and that the only uncertainties are in x , as shown by the vertical error bars. If both x and m are subject to appreciable uncertainties the simplest way to display them is to draw vertical *and* horizontal error bars, whose lengths show the uncertainties in x and m respectively, as in Figure 2.6.

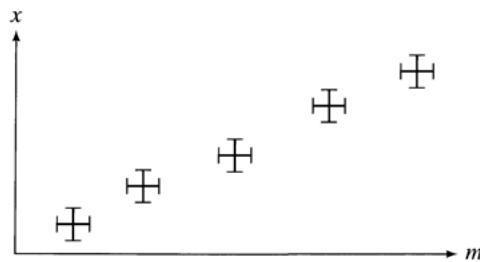


Figure 2.6. Measurements that have uncertainties in both variables can be shown by crosses made up of one error bar for each variable.

Each cross in this plot corresponds to one measurement of x and m , in which x probably lies in the interval defined by the vertical bar of the cross and m probably in that defined by the horizontal bar.

A slightly more complicated possibility is that some quantity may be expected to be proportional to a *power* of another. (For example, the distance traveled by a freely falling object in a time t is $d = \frac{1}{2}gt^2$ and is proportional to the square of t .) Let us suppose that y is expected to be proportional to x^2 . Then

$$y = Ax^2, \quad (2.20)$$

where A is some constant, and a graph of y against x should be a parabola with the general shape of Figure 2.7(a). If we were to measure a series of values for y and x and plot y against x , we might get a graph something like that in Figure 2.7(b). Unfortunately, visually judging whether a set of points such as these fit a parabola (or any other curve, except a straight line) is very hard. A better way to check that $y \propto x^2$ is to plot y against x squared. From Equation (2.20), we see that such a plot should be a straight line, which we can check easily as in Figure 2.7(c).

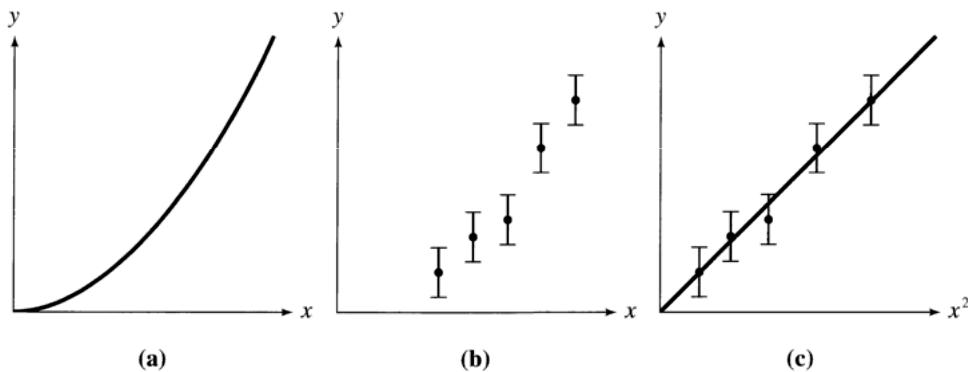


Figure 2.7. (a) If y is proportional to x^2 , a graph of y against x should be a parabola with this general shape. (b) A plot of y against x for a set of measured values is hard to check visually for fit with a parabola. (c) On the other hand, a plot of y against x^2 should be a straight line through the origin, which is easy to check. (In the case shown, we see easily that the points *do* fit a straight line through the origin.)

In the same way, if $y = Ax^n$ (where n is any power), a graph of y against x^n should be a straight line, and by plotting the observed values of y against x^n , we can check easily for such a fit. There are various other situations in which a nonlinear relation (that is, one that gives a curved—nonlinear—graph) can be converted into a linear one by a clever choice of variables to plot. Section 8.6 discusses an important example of such “linearization,” which is worth mentioning briefly here. Often one variable y depends *exponentially* on another variable x :

$$y = Ae^{Bx}.$$

(For example, the activity of a radioactive sample depends exponentially on time.) For such relations, the natural logarithm of y is easily shown to be linear in x ; that is, a graph of $\ln(y)$ against x should be a straight line for an exponential relationship.

Many other, nongraphical ways are available to check the proportionality of two quantities. For example, if $y \propto x$, the ratio y/x should be constant. Thus, having tabulated the measured values of y and x , you could simply add a column to the table that shows the ratios y/x and check that these *ratios* are constant within their experimental uncertainties. Many calculators have a built-in function (called the correlation coefficient) to show how well a set of measurements fits a straight line. (This function is discussed in Section 9.3.) Even when another method is used to check that $y \propto x$, making the graphical check as well is an excellent practice. Graphs such as those in Figures 2.5(b) and (c) show clearly how well (or badly) the measurements verify the predictions; drawing such graphs helps you understand the experiment and the physical laws involved.

2.7 Fractional Uncertainties

The uncertainty δx in a measurement,

$$(\text{measured } x) = x_{\text{best}} \pm \delta x,$$

indicates the reliability or precision of the measurement. The uncertainty δx by itself does not tell the whole story, however. An uncertainty of one inch in a distance of one mile would indicate an unusually precise measurement, whereas an uncertainty of one inch in a distance of three inches would indicate a rather crude estimate. Obviously, the quality of a measurement is indicated not just by the uncertainty δx but also by the *ratio* of δx to x_{best} , which leads us to consider the *fractional uncertainty*,

$$\text{fractional uncertainty} = \frac{\delta x}{|x_{\text{best}}|}. \quad (2.21)$$

(The fractional uncertainty is also called the *relative uncertainty* or the *precision*.) In this definition, the symbol $|x_{\text{best}}|$ denotes the absolute value⁵ of x_{best} . The uncer-

⁵The absolute value $|x|$ of a number x is equal to x when x is positive but is obtained by omitting the minus sign if x is negative. We use the absolute value in (2.21) to guarantee that the fractional uncertainty, like the uncertainty δx itself, is always positive, whether x_{best} is positive or negative. In practice, you can often arrange matters so that measured numbers are positive, and the absolute-value signs in (2.21) can then be omitted.

tainty δx is sometimes called the *absolute uncertainty* to avoid confusion with the fractional uncertainty.

In most serious measurements, the uncertainty δx is much smaller than the measured value x_{best} . Because the fractional uncertainty $\delta x/|x_{\text{best}}|$ is therefore usually a small number, multiplying it by 100 and quoting it as the *percentage uncertainty* is often convenient. For example, the measurement

$$\text{length } l = 50 \pm 1 \text{ cm} \quad (2.22)$$

has a fractional uncertainty

$$\frac{\delta l}{|l_{\text{best}}|} = \frac{1 \text{ cm}}{50 \text{ cm}} = 0.02$$

and a percentage uncertainty of 2%. Thus, the result (2.22) could be given as

$$\text{length } l = 50 \text{ cm} \pm 2\%.$$

Note that although the absolute uncertainty δl has the same units as l , the fractional uncertainty $\delta l/|l_{\text{best}}|$ is a *dimensionless* quantity, without units. Keeping this difference in mind can help you avoid the common mistake of confusing absolute uncertainty with fractional uncertainty.

The fractional uncertainty is an approximate indication of the quality of a measurement, whatever the size of the quantity measured. Fractional uncertainties of 10% or so are usually characteristic of fairly rough measurements. (A rough measurement of 10 inches might have an uncertainty of 1 inch; a rough measurement of 10 miles might have an uncertainty of 1 mile.) Fractional uncertainties of 1 or 2% are characteristic of reasonably careful measurements and are about the best to hope for in many experiments in the introductory physics laboratory. Fractional uncertainties much less than 1% are often hard to achieve and are rather rare in the introductory laboratory.

These divisions are, of course, extremely rough. A few simple measurements can have fractional uncertainties of 0.1% or less with little trouble. A good tape measure can easily measure a distance of 10 feet with an uncertainty of $\frac{1}{10}$ inch, or approximately 0.1%; a good timer can easily measure a period of an hour with an uncertainty of less than a second, or 0.03%. On the other hand, for many quantities that are very hard to measure, a 10% uncertainty would be regarded as an experimental triumph. Large percentage uncertainties, therefore, do not necessarily mean that a measurement is scientifically useless. In fact, many important measurements in the history of physics had experimental uncertainties of 10% or more. Certainly plenty can be learned in the introductory physics laboratory from equipment that has a minimum uncertainty of a few percent.

Quick Check 2.4. Convert the errors in the following measurements of the velocities of two carts on a track into fractional errors and percent errors: (a) $v = 55 \pm 2 \text{ cm/s}$; (b) $u = -20 \pm 2 \text{ cm/s}$. (c) A cart's kinetic energy is measured as $K = 4.58 \text{ J} \pm 2\%$; rewrite this finding in terms of its absolute uncertainty. (Because the uncertainties should be given to one significant figure, you ought to be able to do the calculations in your head.)

2.8 Significant Figures and Fractional Uncertainties

The concept of fractional uncertainty is closely related to the familiar notion of significant figures. In fact, the number of significant figures in a quantity is an approximate indicator of the fractional uncertainty in that quantity. To clarify this connection, let us review briefly the notion of significant figures and recognize that this concept is both approximate and somewhat ambiguous.

To a mathematician, the statement that $x = 21$ to two significant figures means unambiguously that x is closer to 21 than to either 20 or 22; thus, the number 21, with two significant figures, means 21 ± 0.5 . To an experimental scientist, most numbers are numbers that have been read off a meter (or calculated from numbers read off a meter). In particular, if a digital meter displays two significant figures and reads 21, it *may* mean 21 ± 0.5 , but it may also mean 21 ± 1 or even something like 21 ± 5 . (Many meters come with a manual that explains the actual uncertainties.) Under these circumstances, the statement that a measured number has two significant figures is only a rough indicator of its uncertainty. Rather than debate exactly how the concept should be defined, I will adopt a middle-of-the-road definition that 21 with two significant figures means 21 ± 1 , and more generally that a number with N significant figures has an uncertainty of about 1 in the N^{th} digit.

Let us now consider two numbers,

$$x = 21 \quad \text{and} \quad y = 0.21,$$

both of which have been certified accurate to two significant figures. According to the convention just agreed to, these values mean

$$x = 21 \pm 1 \quad \text{and} \quad y = 0.21 \pm 0.01.$$

Although the two numbers both have two significant figures, they obviously have very different uncertainties. On the other hand, they both have the same *fractional uncertainty*, which in this case is 5%:

$$\frac{\delta x}{x} = \frac{\delta y}{y} = \frac{1}{21} = \frac{0.01}{0.21} = 0.05 \text{ or } 5\%.$$

Evidently, the statement that the numbers 21 and 0.21 (or 210, or 2.1, or 0.0021, etc.) have two significant figures is equivalent to saying that they are 5% uncertain. In the same way, 21.0, with three significant figures, is 0.5% uncertain, and so on.

Unfortunately, this useful connection is only approximate. For example, the statement that $s = 10$, with two significant figures, means

$$s = 10 \pm 1 \quad \text{or} \quad 10 \pm 10\%.$$

At the opposite extreme, $t = 99$ (again with two significant figures) means

$$t = 99 \pm 1 \quad \text{or} \quad 99 \pm 1\%.$$

Evidently, the fractional uncertainty associated with two significant figures ranges from 1% to 10%, depending on the first digit of the number concerned.

The approximate correspondence between significant figures and fractional uncertainties can be summarized as in Table 2.4.

Table 2.4. Approximate correspondence between significant figures and fractional uncertainties.

Number of significant figures	Corresponding fractional uncertainty is	
	between	or roughly
1	10% and 100%	50%
2	1% and 10%	5%
3	0.1% and 1%	0.5%

2.9 Multiplying Two Measured Numbers

Perhaps the greatest importance of fractional errors emerges when we start multiplying measured numbers by each other. For example, to find the momentum of a body, we might measure its mass m and its velocity v and then multiply them to give the momentum $p = mv$. Both m and v are subject to uncertainties, which we will have to estimate. The problem, then, is to find the uncertainty in p that results from the known uncertainties in m and v .

First, for convenience, let us rewrite the standard form

$$(\text{measured value of } x) = x_{\text{best}} \pm \delta x$$

in terms of the fractional uncertainty, as

$$(\text{measured value of } x) = x_{\text{best}} \left(1 \pm \frac{\delta x}{|x_{\text{best}}|}\right). \quad (2.23)$$

For example, if the fractional uncertainty is 3%, we see from (2.23) that

$$(\text{measured value of } x) = x_{\text{best}} \left(1 \pm \frac{3}{100}\right);$$

that is, 3% uncertainty means that x probably lies between x_{best} times 0.97 and x_{best} times 1.03,

$$(0.97) \times x_{\text{best}} \leq x \leq (1.03) \times x_{\text{best}}.$$

We will find this a useful way to think about a measured number that we will have to multiply.

Let us now return to our problem of calculating $p = mv$, when m and v have been measured, as

$$(\text{measured } m) = m_{\text{best}} \left(1 \pm \frac{\delta m}{|m_{\text{best}}|}\right) \quad (2.24)$$

and

$$(\text{measured } v) = v_{\text{best}} \left(1 \pm \frac{\delta v}{|v_{\text{best}}|}\right) \quad (2.25)$$

Because m_{best} and v_{best} are our best estimates for m and v , our best estimate for $p = mv$ is

$$(\text{best estimate for } p) = p_{\text{best}} = m_{\text{best}}v_{\text{best}}.$$

The largest probable values of m and v are given by (2.24) and (2.25) with the plus signs. Thus, the largest probable value for $p = mv$ is

$$(\text{largest value for } p) = m_{\text{best}}v_{\text{best}} \left(1 + \frac{\delta m}{|m_{\text{best}}|}\right) \left(1 + \frac{\delta v}{|v_{\text{best}}|}\right). \quad (2.26)$$

The smallest probable value for p is given by a similar expression with two minus signs. Now, the product of the parentheses in (2.26) can be multiplied out as

$$\left(1 + \frac{\delta m}{|m_{\text{best}}|}\right) \left(1 + \frac{\delta v}{|v_{\text{best}}|}\right) = 1 + \frac{\delta m}{|m_{\text{best}}|} + \frac{\delta v}{|v_{\text{best}}|} + \frac{\delta m}{|m_{\text{best}}|} \frac{\delta v}{|v_{\text{best}}|}. \quad (2.27)$$

Because the two fractional uncertainties $\delta m/|m_{\text{best}}|$ and $\delta v/|v_{\text{best}}|$ are small numbers (a few percent, perhaps), their product is extremely small. Therefore, the last term in (2.27) can be neglected. Returning to (2.26), we find

$$(\text{largest value of } p) = m_{\text{best}}v_{\text{best}} \left(1 + \frac{\delta m}{|m_{\text{best}}|} + \frac{\delta v}{|v_{\text{best}}|}\right).$$

The smallest probable value is given by a similar expression with two minus signs. Our measurements of m and v , therefore, lead to a value of $p = mv$ given by

$$(\text{value of } p) = m_{\text{best}}v_{\text{best}} \left(1 \pm \left[\frac{\delta m}{|m_{\text{best}}|} + \frac{\delta v}{|v_{\text{best}}|}\right]\right).$$

Comparing this equation with the general form

$$(\text{value of } p) = p_{\text{best}} \left(1 \pm \frac{\delta p}{|p_{\text{best}}|}\right),$$

we see that the best estimate for p is $p_{\text{best}} = m_{\text{best}}v_{\text{best}}$ (as we already knew) and that the *fractional uncertainty in p is the sum of the fractional uncertainties in m and v* ,

$$\frac{\delta p}{|p_{\text{best}}|} = \frac{\delta m}{|m_{\text{best}}|} + \frac{\delta v}{|v_{\text{best}}|}.$$

If, for example, we had the following measurements for m and v ,

$$m = 0.53 \pm 0.01 \text{ kg}$$

and

$$v = 9.1 \pm 0.3 \text{ m/s},$$

the best estimate for $p = mv$ is

$$p_{\text{best}} = m_{\text{best}}v_{\text{best}} = (0.53) \times (9.1) = 4.82 \text{ kg}\cdot\text{m/s}.$$

To compute the uncertainty in p , we would first compute the fractional errors

$$\frac{\delta m}{m_{\text{best}}} = \frac{0.01}{0.53} = 0.02 = 2\%$$

and

$$\frac{\delta v}{v_{\text{best}}} = \frac{0.3}{9.1} = 0.03 = 3\%.$$

The fractional uncertainty in p is then the sum:

$$\frac{\delta p}{p_{\text{best}}} = 2\% + 3\% = 5\%.$$

If we want to know the absolute uncertainty in p , we must multiply by p_{best} :

$$\delta p = \frac{\delta p}{p_{\text{best}}} \times p_{\text{best}} = 0.05 \times 4.82 = 0.241.$$

We then round δp and p_{best} to give us our final answer

$$(\text{value of } p) = 4.8 \pm 0.2 \text{ kg}\cdot\text{m/s.}$$

The preceding considerations apply to any product of two measured quantities. We have therefore discovered our second general rule for the propagation of errors. If we measure any two quantities x and y and form their product, the uncertainties in the original two quantities “propagate” to cause an uncertainty in their product. This uncertainty is given by the following rule:

**Uncertainty in a Product
(Provisional Rule)**

If two quantities x and y have been measured with small fractional uncertainties $\delta x/|x_{\text{best}}|$ and $\delta y/|y_{\text{best}}|$, and if the measured values of x and y are used to calculate the product $q = xy$, then the *fractional uncertainty in q is the sum of the fractional uncertainties in x and y* ,

$$\frac{\delta q}{|q_{\text{best}}|} \approx \frac{\delta x}{|x_{\text{best}}|} + \frac{\delta y}{|y_{\text{best}}|}. \quad (2.28)$$

I call this rule “provisional,” because, just as with the rule for uncertainty in a difference, I will replace it with a more precise rule later on. Two other features of this rule also need to be emphasized. First, the derivation of (2.28) required that the fractional uncertainties in x and y both be small enough that we could neglect their product. This requirement is almost always true in practice, and I will always assume it. Nevertheless, remember that if the fractional uncertainties are *not* much smaller than 1, the rule (2.28) may not apply. Second, even when x and y have different dimensions, (2.28) balances dimensionally because all fractional uncertainties are dimensionless.

In physics, we frequently multiply numbers together, and the rule (2.28) for finding the uncertainty in a product will obviously be an important tool in error analysis. For the moment, our main purpose is to emphasize that the uncertainty in any product $q = xy$ is expressed most simply in terms of fractional uncertainties, as in (2.28).

Quick Check 2.5. To find the area of a rectangular plate, a student measures its sides as $l = 9.1 \pm 0.1$ cm and $b = 3.3 \pm 0.1$ cm. Express these uncertainties as percent uncertainties and then find the student's answer for the area $A = lb$ with its uncertainty. (Find the latter as a percent uncertainty first and then convert to an absolute uncertainty. Do all error calculations in your head.)

Principal Definitions and Equations of Chapter 2

STANDARD FORM FOR STATING UNCERTAINTIES

The standard form for reporting a measurement of a physical quantity x is

$$(\text{measured value of } x) = x_{\text{best}} \pm \delta x,$$

where

$$x_{\text{best}} = (\text{best estimate for } x)$$

and

$$\delta x = (\text{uncertainty or error in the measurement}). \quad [\text{See (2.3)}]$$

This statement expresses our confidence that the correct value of x probably lies in (or close to) the range from $x_{\text{best}} - \delta x$ to $x_{\text{best}} + \delta x$.

DISCREPANCY

The *discrepancy* between two measured values of the same physical quantity is

$$\text{discrepancy} = \text{difference between two measured} \\ \text{values of the same quantity.} \quad [\text{See (2.10)}]$$

FRACTIONAL UNCERTAINTY

If x is measured in the standard form $x_{\text{best}} \pm \delta x$, the *fractional uncertainty* in x is

$$\text{fractional uncertainty} = \frac{\delta x}{|x_{\text{best}}|}. \quad [\text{See (2.21)}]$$

The *percent uncertainty* is just the fractional uncertainty expressed as a percentage (that is, multiplied by 100%).

We have found two provisional rules, (2.18) and (2.28), for error propagation that show how the uncertainties in two quantities x and y propagate to cause uncertainties in calculations of the difference $x - y$ or the product xy . A complete discussion of error propagation appears in Chapter 3, where I show that the rules (2.18) and (2.28) can frequently be replaced with more refined rules (given in Section 3.6). For this reason, I have not reproduced (2.18) and (2.28) here.

Problems for Chapter 2

Notes: The problems at the end of each chapter are arranged by section number. A problem listed for a specific section may, of course, involve ideas from previous sections but does not require knowledge of later sections. Therefore, you may try problems listed for a specific section as soon as you have read that section.

The approximate difficulty of each problem is indicated by one, two, or three stars. A one-star problem should be straightforward and usually involves a single concept. Two-star problems are more difficult or require more work (drawing a graph, for instance). Three-star problems are the most difficult and may require considerably more labor.

Answers to the odd-numbered problems can be found in the Answers Section at the back of the book.

For Section 2.1: Best Estimate \pm Uncertainty

2.1. ★ In Chapter 1, a carpenter reported his measurement of the height of a doorway by stating that his best estimate was 210 cm and that he was confident the height was between 205 and 215 cm. Rewrite this result in the standard form $x_{\text{best}} \pm \delta x$. Do the same for the measurements reported in Equations (1.1), (1.2), and (1.4).

2.2. ★ A student studying the motion of a cart on an air track measures its position, velocity, and acceleration at one instant, with the results shown in Table 2.5. Rewrite these results in the standard form $x_{\text{best}} \pm \delta x$.

Table 2.5. Measurements of position, velocity, and acceleration; for Problem 2.2.

Variable	Best estimate	Probable range
Position, x	53.3	53.1 to 53.5 (cm)
Velocity, v	-13.5	-14.0 to -13.0 (cm/s)
Acceleration, a	93	90 to 96 (cm/s ²)

For Section 2.2: Significant Figures

2.3. ★ Rewrite the following results in their clearest forms, with suitable numbers of significant figures:

- (a) measured height = 5.03 ± 0.04329 m
- (b) measured time = 1.5432 ± 1 s
- (c) measured charge = $-3.21 \times 10^{-19} \pm 2.67 \times 10^{-20}$ C
- (d) measured wavelength = $0.000,000,563 \pm 0.000,000,07$ m
- (e) measured momentum = $3.267 \times 10^3 \pm 42$ g·cm/s.

2.4. ★ Rewrite the following equations in their clearest and most appropriate forms:

- (a) $x = 3.323 \pm 1.4$ mm
- (b) $t = 1,234,567 \pm 54,321$ s
- (c) $\lambda = 5.33 \times 10^{-7} \pm 3.21 \times 10^{-9}$ m
- (d) $r = 0.000,000,538 \pm 0.000,000,03$ mm

For Section 2.3: Discrepancy

2.5. ★ Two students measure the length of the same rod and report the results 135 ± 3 mm and 137 ± 3 mm. Draw an illustration like that in Figure 2.1 to represent these two measurements. What is the discrepancy between the two measurements, and is it significant?

2.6. ★ Each of two research groups discovers a new elementary particle. The two reported masses are

$$m_1 = (7.8 \pm 0.1) \times 10^{-27} \text{ kg}$$

and

$$m_2 = (7.0 \pm 0.2) \times 10^{-27} \text{ kg.}$$

Draw an illustration like that in Figure 2.1 to represent these two measurements. The question arises whether these two measurements could actually be of the same particle. Based on the reported masses, would you say they are likely to be the same particle? In particular, what is the discrepancy in the two measurements (assuming they really are measurements of the same mass)?

For Section 2.4: Comparison of Measured and Accepted Values

2.7. ★ (a) A student measures the density of a liquid five times and gets the results (all in gram/cm³) 1.8, 2.0, 2.0, 1.9, and 1.8. What would you suggest as the best estimate and uncertainty based on these measurements? **(b)** The student is told that the accepted value is 1.85 gram/cm³. What is the discrepancy between the student's best estimate and the accepted value? Do you think it is significant?

2.8. ★ Two groups of students measure the charge of the electron and report their results as follows:

$$\text{Group A: } e = (1.75 \pm 0.04) \times 10^{-19} \text{ C}$$

and

$$\text{Group B: } e = (1.62 \pm 0.04) \times 10^{-19} \text{ C.}$$

What should each group report for the discrepancy between its value and the accepted value,

$$e = 1.60 \times 10^{-19} \text{ C}$$

(with negligible uncertainty)? Draw an illustration similar to that in Figure 2.2 to show these results and the accepted value. Which of the results would you say is satisfactory?

For Section 2.5: Comparison of Two Measured Numbers

2.9. ★ In an experiment on the simple pendulum, a student uses a steel ball suspended from a light string, as shown in Figure 2.8. The effective length l of the

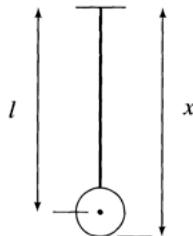


Figure 2.8. A simple pendulum; for Problem 2.9.

pendulum is the distance from the top of the string to the *center* of the ball, as shown. To find l , he first measures the distance x from the top of the string to the bottom of the ball and the radius r of the ball; he then subtracts to give $l = x - r$. If his two measurements are

$$x = 95.8 \pm 0.1 \text{ cm} \quad \text{and} \quad r = 2.30 \pm 0.02 \text{ cm},$$

what should be his answer for the length l and its uncertainty, as given by the provisional rule (2.18)?

2.10. ★ The time a carousel takes to make one revolution is measured by noting the starting and stopping times using the second hand of a wrist watch and subtracting. If the starting and stopping times are uncertain by ± 1 second each, what is the uncertainty in the time for one revolution, as given by the provisional rule (2.18)?

2.11. ★ In an experiment to check conservation of angular momentum, a student obtains the results shown in Table 2.6 for the initial and final angular momenta (L and L') of a rotating system. Add an extra column to the table to show the difference $L - L'$ and its uncertainty. Are the student's results consistent with conservation of angular momentum?

Table 2.6. Initial and final angular momenta (in $\text{kg}\cdot\text{m}^2/\text{s}$); for Problems 2.11 and 2.14.

Initial L	Final L'
3.0 ± 0.3	2.7 ± 0.6
7.4 ± 0.5	8.0 ± 1
14.3 ± 1	16.5 ± 1
25 ± 2	24 ± 2
32 ± 2	31 ± 2
37 ± 2	41 ± 2

2.12. ★ The acceleration a of a cart sliding down a frictionless incline with slope θ is expected to be $g\sin\theta$. To test this, a student measures the acceleration a of a cart on an incline for several different values of θ ; she also calculates the corresponding expected accelerations $g\sin\theta$ for each θ and obtains the results shown in Table 2.7. Add a column to the table to show the discrepancies $a - g\sin\theta$ and their uncertainties. Do the results confirm that a is given by $g\sin\theta$? If not, can you suggest a reason they do not?

Table 2.7. Measured and expected accelerations; for Problem 2.12.

Trial number	Acceleration a (m/s^2)	Expected acceleration $g\sin\theta$ (m/s^2)
1	2.04 ± 0.04	2.36 ± 0.1
2	3.58 ± 0.06	3.88 ± 0.08
3	4.32 ± 0.08	4.57 ± 0.05
4	4.85 ± 0.09	5.05 ± 0.04
5	5.53 ± 0.1	5.72 ± 0.03

2.13. ★★ An experimenter measures the separate masses M and m of a car and trailer. He gives his results in the standard form $M_{\text{best}} \pm \delta M$ and $m_{\text{best}} \pm \delta m$. What would be his best estimate for the total mass $M + m$? By considering the largest and smallest probable values of the total mass, show that his uncertainty in the total mass is just the sum of δM and δm . State your arguments clearly; don't just write down the answer. (This problem provides another example of error propagation: The uncertainties in the measured numbers, M and m , propagate to cause an uncertainty in the sum $M + m$.)

For Section 2.6: Checking Relationships with a Graph

2.14. ★★ Using the data of Problem 2.11, make a plot of final angular momentum L' against initial angular momentum L for the experiment described there. (Include

vertical and horizontal error bars, and be sure to include the origin. As with all graphs, label your axes, including units, use squared paper, and choose the scales so that the graph fills a good proportion of the page.) On what curve would you expect the points to lie? Do they lie on this curve within experimental uncertainties?

2.15. ★★ According to the ideal gas law, if the volume of a gas is kept constant, the pressure P should be proportional to the absolute temperature T . To check this proportionality, a student measures the pressure of a gas at five different temperatures (always with the same volume) and gets the results shown in Table 2.8. Plot these results in a graph of P against T , and decide whether they confirm the expected proportionality of P and T .

Table 2.8. Temperature and pressure of a gas; for Problem 2.15.

Temperature (K) (negligible uncertainty)	Pressure (atm) (all ± 0.04)
100	0.36
150	0.46
200	0.71
250	0.83
300	1.04

2.16. ★★ You have learned (or will learn) in optics that certain lenses (namely, thin spherical lenses) can be characterized by a parameter called the focal length f and that if an object is placed at a distance p from the lens, the lens forms an image at a distance q , satisfying the *lens equation*, $1/f = (1/p) + (1/q)$, where f always has the same value for a given lens. To check if these ideas apply to a certain lens, a student places a small light bulb at various distances p from the lens and measures the location q of the corresponding images. She then calculates the corresponding values of f from the lens equation and obtains the results shown in Table 2.9. Make a plot of f against p , with appropriate error bars, and decide if it is true that this particular lens has a unique focal length f .

Table 2.9. Object distances p (in cm) and corresponding focal lengths f (in cm); for Problem 2.16.

Object distance p (negligible uncertainty)	Focal length f (all ± 2)
45	28
55	34
65	33
75	37
85	40

2.17. ★★ The power P delivered to a resistance R by a current I is supposed to be given by the relation $P = RI^2$. To check this relation, a student sends several different currents through an unknown resistance immersed in a cup of water and measures the power delivered (by measuring the water's rise in temperature). Use the results shown in Table 2.10 to make plots of P against I and P against I^2 , including error bars. Use the second plot to decide if this experiment is consistent with the expected proportionality of P and I^2 .

Table 2.10. Current I and power P ; for Problem 2.17.

Current I (amps) (negligible uncertainty)	Power P (watts) (all ± 50)
1.5	270
2.0	380
2.5	620
3.0	830
3.5	1280
4.0	1600

2.18. ★★★ If a stone is thrown vertically upward with speed v , it should rise to a height h given by $v^2 = 2gh$. In particular, v^2 should be proportional to h . To test this proportionality, a student measures v^2 and h for seven different throws and gets the results shown in Table 2.11. **(a)** Make a plot of v^2 against h , including vertical and horizontal error bars. (As usual, use squared paper, label your axes, and choose your scale sensibly.) Is your plot consistent with the prediction that $v^2 \propto h$? **(b)** The slope of your graph should be $2g$. To find the slope, draw what seems to be the best straight line through the points and then measure its slope. To find the uncertainty in the slope, draw the steepest and least steep lines that seem to fit the data reasonably. The slopes of these lines give the largest and smallest probable values of the slope. Are your results consistent with the accepted value $2g = 19.6 \text{ m/s}^2$?

Table 2.11. Heights and speeds of a stone thrown vertically upward; for Problem 2.18.

h (m) all ± 0.05	v^2 (m^2/s^2)
0.4	7 ± 3
0.8	17 ± 3
1.4	25 ± 3
2.0	38 ± 4
2.6	45 ± 5
3.4	62 ± 5
3.8	72 ± 6

2.19. ★★★ In an experiment with a simple pendulum, a student decides to check whether the period T is independent of the amplitude A (defined as the largest angle that the pendulum makes with the vertical during its oscillations). He obtains the

Table 2.12. Amplitude and period of a pendulum; for Problem 2.19.

Amplitude A (deg)	Period T (s)
5 ± 2	1.932 ± 0.005
17 ± 2	1.94 ± 0.01
25 ± 2	1.96 ± 0.01
40 ± 4	2.01 ± 0.01
53 ± 4	2.04 ± 0.01
67 ± 6	2.12 ± 0.02

results shown in Table 2.12. (a) Draw a graph of T against A . (Consider your choice of scales carefully. If you have any doubt about this choice, draw two graphs, one including the origin, $A = T = 0$, and one in which only values of T between 1.9 and 2.2 s are shown.) Should the student conclude that the period is independent of the amplitude? (b) Discuss how the conclusions of part (a) would be affected if all the measured values of T had been uncertain by ± 0.3 s.

For Section 2.7: Fractional Uncertainties

2.20. ★ Compute the percentage uncertainties for the five measurements reported in Problem 2.3. (Remember to round to a reasonable number of significant figures.)

2.21. ★ Compute the percentage uncertainties for the four measurements in Problem 2.4.

2.22. ★ Convert the percent errors given for the following measurements into absolute uncertainties and rewrite the results in the standard form $x_{\text{best}} \pm \delta x$ rounded appropriately.

- (a) $x = 543.2 \text{ m} \pm 4\%$
- (b) $v = -65.9 \text{ m/s} \pm 8\%$
- (c) $\lambda = 671 \times 10^{-9} \text{ m} \pm 4\%$

2.23. ★ A meter stick can be read to the nearest millimeter; a traveling microscope can be read to the nearest 0.1 mm. Suppose you want to measure a length of 2 cm with a precision of 1%. Can you do so with the meter stick? Is it possible to do so with the microscope?

2.24. ★ (a) A digital voltmeter reads voltages to the nearest thousandth of a volt. What will be its percent uncertainty in measuring a voltage of approximately 3 volts? (b) A digital balance reads masses to the nearest hundredth of a gram. What will be its percent uncertainty in measuring a mass of approximately 6 grams?

2.25. ★★ To find the acceleration of a cart, a student measures its initial and final velocities, v_i and v_f , and computes the difference $(v_f - v_i)$. Her data in two separate

Table 2.13. Initial and final velocities (all in cm/s and all $\pm 1\%$); for Problem 2.25.

	v_i	v_f
First run	14.0	18.0
Second run	19.0	19.6

trials are shown in Table 2.13. All have an uncertainty of $\pm 1\%$. (a) Calculate the absolute uncertainties in all four measurements; find the change ($v_f - v_i$) and its uncertainty in each run. (b) Compute the percent uncertainty for each of the two values of ($v_f - v_i$). Your answers, especially for the second run, illustrate the disastrous results of finding a small number by taking the difference of two much larger numbers.

For Section 2.8: Significant Figures and Fractional Uncertainties

2.26. ★ (a) A student's calculator shows an answer 123.123. If the student decides that this number actually has only three significant figures, what are its absolute and fractional uncertainties? (To be definite, adopt the convention that a number with N significant figures is uncertain by ± 1 in the N^{th} digit.) **(b)** Do the same for the number 1231.23. **(c)** Do the same for the number 321.321. **(d)** Do the fractional uncertainties lie in the range expected for three significant figures?

2.27. ★★ (a) My calculator gives the answer $x = 6.1234$, but I know that x has a fractional uncertainty of 2%. Restate my answer in the standard form $x_{\text{best}} \pm \delta x$ properly rounded. How many significant figures does the answer really have? **(b)** Do the same for $y = 1.1234$ with a fractional uncertainty of 2%. **(c)** Likewise, for $z = 9.1234$.

For Section 2.9: Multiplying Two Measured Numbers

2.28. ★ (a) A student measures two quantities a and b and obtains the results $a = 11.5 \pm 0.2 \text{ cm}$ and $b = 25.4 \pm 0.2 \text{ s}$. She now calculates the product $q = ab$. Find her answer, giving both its percent and absolute uncertainties, as found using the provisional rule (2.28). **(b)** Repeat part (a) using $a = 5.0 \text{ m} \pm 7\%$ and $b = 3.0 \text{ N} \pm 1\%$.

2.29. ★ (a) A student measures two quantities a and b and obtains the results $a = 10 \pm 1 \text{ N}$ and $b = 272 \pm 1 \text{ s}$. He now calculates the product $q = ab$. Find his answer, giving both its percent and absolute uncertainties, as found using the provisional rule (2.28). **(b)** Repeat part (a) using $a = 3.0 \text{ ft} \pm 8\%$ and $b = 4.0 \text{ lb} \pm 2\%$.

2.30. ★★ A well-known rule states that when two numbers are multiplied together, the answer will be reliable if rounded to the number of significant figures in the less precise of the original two numbers. **(a)** Using our rule (2.28) and the fact that significant figures correspond roughly to fractional uncertainties, prove that this rule

is *approximately* valid. (To be definite, treat the case that the less precise number has two significant figures.) **(b)** Show by example that the answer can actually be somewhat less precise than the “well-known” rule suggests. (This reduced precision is especially true if several numbers are multiplied together.)

2.31. ★★ (a) A student measures two numbers x and y as

$$x = 10 \pm 1 \quad \text{and} \quad y = 20 \pm 1.$$

What is her best estimate for their product $q = xy$? Using the largest probable values for x and y (11 and 21), calculate the largest probable value of q . Similarly, find the smallest probable value of q , and hence the range in which q probably lies. Compare your result with that given by the rule (2.28). **(b)** Do the same for the measurements

$$x = 10 \pm 8 \quad \text{and} \quad y = 20 \pm 15.$$

[Remember that the rule (2.28) was derived by assuming that the fractional uncertainties are much less than 1.]

Chapter 3

Propagation of Uncertainties

Most physical quantities usually cannot be measured in a single direct measurement but are instead found in two distinct steps. First, we measure one or more quantities that *can* be measured directly and from which the quantity of interest can be calculated. Second, we use the measured values of these quantities to calculate the quantity of interest itself. For example, to find the area of a rectangle, you actually measure its length l and height h and then calculate its area A as $A = lh$. Similarly, the most obvious way to find the velocity v of an object is to measure the distance traveled, d , and the time taken, t , and then to calculate v as $v = d/t$. Any reader with experience in an introductory laboratory can easily think of more examples. In fact, a little thought will show that almost all interesting measurements involve these two distinct steps of direct measurement followed by calculation.

When a measurement involves these two steps, the estimation of uncertainties also involves two steps. We must first estimate the uncertainties in the quantities measured directly and then determine how these uncertainties “propagate” through the calculations to produce an uncertainty in the final answer.¹ This propagation of errors is the main subject of this chapter.

In fact, examples of propagation of errors were presented in Chapter 2. In Section 2.5, I discussed what happens when two numbers x and y are measured and the results are used to calculate the difference $q = x - y$. We found that the uncertainty in q is just the *sum* $\delta q \approx \delta x + \delta y$ of the uncertainties in x and y . Section 2.9 discussed the product $q = xy$, and Problem 2.13 discussed the sum $q = x + y$. I review these cases in Section 3.3; the rest of this chapter is devoted to more general cases of propagation of uncertainties and includes several examples.

Before I address error propagation in Section 3.3, I will briefly discuss the estimation of uncertainties in quantities measured directly in Sections 3.1 and 3.2. The methods presented in Chapter 1 are reviewed, and further examples are given of error estimation in direct measurements.

Starting in Section 3.3, I will take up the propagation of errors. You will learn that almost all problems in error propagation can be solved using three simple rules.

¹In Chapter 4, I discuss another way in which the final uncertainty can sometimes be estimated. If all measurements can be repeated several times, and if all uncertainties are known to be random in character, then the uncertainty in the quantity of interest can be estimated by examining the spread in answers. Even when this method is possible, it is usually best used as a check on the two-step procedure discussed in this chapter.

A single, more complicated, rule will also be presented that covers all cases and from which the three simpler rules can be derived.

This chapter is long, but its length simply reflects its great importance. Error propagation is a technique you will use repeatedly in the laboratory, and you need to become familiar with the methods described here. The only exception is that the material of Section 3.11 is not used again until Section 5.6; thus, if the ideas of this chapter are all new to you, consider skipping Section 3.11 on your first reading.

3.1 Uncertainties in Direct Measurements

Almost all direct measurements involve reading a scale (on a ruler, clock, or voltmeter, for example) or a digital display (on a digital clock or voltmeter, for example). Some problems in scale reading were discussed in Section 1.5. Sometimes the main sources of uncertainty are the reading of the scale and the need to interpolate between the scale markings. In such situations, a reasonable estimate of the uncertainty is easily made. For example, if you have to measure a clearly defined length l with a ruler graduated in millimeters, you might reasonably decide that the length could be read to the nearest millimeter but no better. Here, the uncertainty δl would be $\delta l = 0.5 \text{ mm}$. If the scale markings are farther apart (as with tenths of an inch), you might reasonably decide you could read to one-fifth of a division, for example. In any case, the uncertainties associated with the reading of a scale can obviously be estimated quite easily and realistically.

Unfortunately, other sources of uncertainty are frequently much more important than difficulties in scale reading. In measuring the distance between two points, your main problem may be to decide where those two points really are. For example, in an optics experiment, you may wish to measure the distance q from the center of a lens to a focused image, as in Figure 3.1. In practice, the lens is usually several millimeters thick, so locating its center is hard; if the lens comes in a bulky mounting, as it often does, locating the center is even harder. Furthermore, the image may appear to be well-focused throughout a range of many millimeters. Even though the apparatus is mounted on an optical bench that is clearly graduated in millimeters, the uncertainty in the distance from lens to image could easily be a centimeter or so. Since this uncertainty arises because the two points concerned are not clearly defined, this kind of problem is called a *problem of definition*.

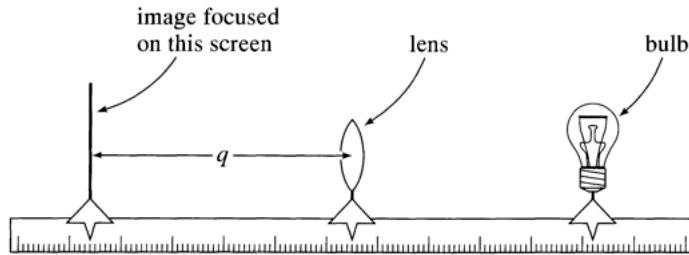


Figure 3.1. An image of the light bulb on the right is focused by the lens onto the screen at the left.

This example illustrates a serious danger in error estimation. If you look only at the scales and forget about other sources of uncertainty, you can badly underestimate the total uncertainty. In fact, the beginning student's most common mistake is to overlook some sources of uncertainty and hence *underestimate* uncertainties, often by a factor of 10 or more. Of course, you must also avoid *overestimating* errors. Experimenters who decide to play safe and to quote generous uncertainties on all measurements may avoid embarrassing inconsistencies, but their measurements may not be of much use. Clearly, the ideal is to find all possible causes of uncertainty and estimate their effects accurately, which is often not quite as hard as it sounds.

Superficially, at least, reading a digital meter is much easier than a conventional analog meter. Unless a digital meter is defective, it should display only significant figures. Thus, it is usually safe to say that the number of significant figures in a digital reading is precisely the number of figures displayed. Unfortunately, as discussed in Section 2.8, the exact meaning of significant figures is not always clear. Thus, a digital voltmeter that tells us that $V = 81$ microvolts could mean that the uncertainty is anything from $\delta V = 0.5$ to $\delta V = 1$ or more. Without a manual to tell you the uncertainty in a digital meter, a reasonable assumption is that the uncertainty in the final digit is ± 1 (so that the voltage just mentioned is $V = 81 \pm 1$).

The digital meter, even more than the analog scale, can give a misleading impression of accuracy. For example, a student might use a digital timer to time the fall of a weight in an Atwood machine or similar device. If the timer displays 8.01 seconds, the time of fall is apparently

$$t = 8.01 \pm 0.01 \text{ s.} \quad (3.1)$$

However, the careful student who repeats the experiment under nearly identical conditions might find a second measurement of 8.41 s; that is,

$$t = 8.41 \pm 0.01 \text{ s.}$$

One likely explanation of this large discrepancy is that uncertainties in the starting procedure vary the initial conditions and hence the time of fall; that is, the measured times really are different. In any case, the accuracy claimed in Equation (3.1) clearly is ridiculously too good. Based on the two measurements made, a more realistic answer would be

$$t = 8.2 \pm 0.2 \text{ s.}$$

In particular, the uncertainty is some 20 times larger than suggested in Equation (3.1) based on the original single reading.

This example brings us to another point mentioned in Chapter 1: Whenever a measurement can be repeated, it should usually be made several times. The resulting spread of values often provides a good indication of the uncertainties, and the average of the values is almost certainly more trustworthy than any one measurement. Chapters 4 and 5 discuss the statistical treatment of multiple measurements. Here, I emphasize only that if a measurement is repeatable, it should be repeated, both to obtain a more reliable answer (by averaging) and, more important, to get an estimate of the uncertainties. Unfortunately, as also mentioned in Chapter 1, repeating a measurement does not always reveal uncertainties. If the measurement is subject to a systematic error, which pushes all results in the same direction (such as a clock that

runs slow), the spread in results will not reflect this systematic error. Eliminating such systematic errors requires careful checks of calibration and procedures.

3.2 The Square-Root Rule for a Counting Experiment

Another, different kind of direct measurement has an uncertainty that can be estimated easily. Some experiments require you to count events that occur at random but have a definite average rate. For example, the babies born in a hospital arrive in a fairly random way, but in the long run births in any one hospital probably occur at a definite average rate. Imagine that a demographer who wants to know this rate counts 14 births in a certain two-week period at a local hospital. Based on this result, he would naturally say that his best estimate for the expected number of births in two weeks is 14. Unless he has made a mistake, 14 is *exactly* the number of births in the two-week period he chose to observe. Because of the random way births occur, however, 14 obviously may *not* equal the actual average number of births in all two-week periods. Perhaps this number is 13, 15, or even a fractional number such as 13.5 or 14.7.

Evidently, the uncertainty in this kind of experiment is not in the observed number counted (14 in our example). Instead, the uncertainty is in how well this observed number approximates the true average number. The problem is to estimate how large this uncertainty is. Although I discuss the theory of these counting experiments in Chapter 11, the answer is remarkably simple and is easily stated here: The uncertainty in any counted number of random events, as an estimate of the true average number, is *the square root of the counted number*. In our example, the demographer counted 14 births in a certain two-week period. Therefore, his uncertainty is $\sqrt{14} \approx 4$, and his final conclusion would be

$$\text{(average births in a two-week period)} = 14 \pm 4.$$

To make this statement more general, suppose we count the occurrences of any event (such as the births of babies in a hospital) that occurs randomly but at a definite average rate. Suppose we count for a chosen time interval T (such as two weeks), and we denote the number of observed events by the Greek letter ν . (Pronounced “nu,” this symbol is the Greek form of the letter n and stands for *number*.) Based on this experiment, our best estimate for the average number of events in time T is, of course, the observed number ν , and the uncertainty in this estimate is the square root of the number, that is, $\sqrt{\nu}$. Therefore, our answer for the average number of events in time T is

$$\boxed{\text{(average number of events in time } T\text{)} = \nu \pm \sqrt{\nu}.} \quad (3.2)$$

I refer to this important result as the *Square-Root Rule for Counting Experiments*.

Counting experiments of this type occur frequently in the physics laboratory. The most prominent example is in the study of radioactivity. In a radioactive material, each nucleus decays at a random time, but the decays in a large sample occur at a definite average rate. To find this rate, you can simply count the number ν of

decays in some convenient time interval T ; the expected number of decays in time T , with its uncertainty, is then given by the square-root rule, (3.2).

Quick Check 3.1. (a) To check the activity of a radioactive sample, an inspector places the sample in a liquid scintillation counter to count the number of decays in a two-minute interval and obtains 33 counts. What should he report as the number of decays produced by the sample in two minutes? (b) Suppose, instead, he had monitored the same sample for 50 minutes and obtained 907 counts. What would be his answer for the number of decays in 50 minutes? (c) Find the percent uncertainties in these two measurements, and comment on the usefulness of counting for a longer period as in part (b).

3.3 Sums and Differences; Products and Quotients

For the remainder of this chapter, I will suppose that we have measured one or more quantities x, y, \dots , with corresponding uncertainties $\delta x, \delta y, \dots$, and that we now wish to use the measured values of x, y, \dots , to calculate the quantity of real interest, q . The calculation of q is usually straightforward; the problem is how the uncertainties, $\delta x, \delta y, \dots$, propagate through the calculation and lead to an uncertainty δq in the final value of q .

SUMS AND DIFFERENCES

Chapter 2 discussed what happens when you measure two quantities x and y and calculate their sum, $x + y$, or their difference, $x - y$. To estimate the uncertainty in the sum or difference, we had only to decide on their highest and lowest probable values. The highest and lowest probable values of x are $x_{\text{best}} \pm \delta x$, and those of y are $y_{\text{best}} \pm \delta y$. Hence, the highest probable value of $x + y$ is

$$x_{\text{best}} + y_{\text{best}} + (\delta x + \delta y),$$

and the lowest probable value is

$$x_{\text{best}} + y_{\text{best}} - (\delta x + \delta y).$$

Thus, the best estimate for $q = x + y$ is

$$q_{\text{best}} = x_{\text{best}} + y_{\text{best}},$$

and its uncertainty is

$$\delta q \approx \delta x + \delta y. \quad (3.3)$$

A similar argument (be sure you can reconstruct it) shows that the uncertainty in the difference $x - y$ is given by the same formula (3.3). That is, the uncertainty in either the sum $x + y$ or the difference $x - y$ is the sum $\delta x + \delta y$ of the uncertainties in x and y .

If we have several numbers x, \dots, w to be added or subtracted, then repeated application of (3.3) gives the following provisional rule.

Uncertainty in Sums and Differences (Provisional Rule)

If several quantities x, \dots, w are measured with uncertainties $\delta x, \dots, \delta w$, and the measured values used to compute

$$q = x + \dots + z - (u + \dots + w),$$

then the uncertainty in the computed value of q is the sum,

$$\delta q \approx \delta x + \dots + \delta z + \delta u + \dots + \delta w, \quad (3.4)$$

of all the original uncertainties.

In other words, when you add or subtract any number of quantities, the uncertainties in those quantities always *add*. As before, I use the sign \approx to emphasize that this rule is only provisional.

Example: Adding and Subtracting Masses

As a simple example of rule (3.4), suppose an experimenter mixes together the liquids in two flasks, having first measured their separate masses when full and empty, as follows:

$$\begin{aligned} M_1 &= \text{mass of first flask and contents} &= 540 \pm 10 \text{ grams} \\ m_1 &= \text{mass of first flask empty} &= 72 \pm 1 \text{ grams} \\ M_2 &= \text{mass of second flask and contents} &= 940 \pm 20 \text{ grams} \\ m_2 &= \text{mass of second flask empty} &= 97 \pm 1 \text{ grams} \end{aligned}$$

He now calculates the total mass of liquid as

$$\begin{aligned} M &= M_1 - m_1 + M_2 - m_2 \\ &= (540 - 72 + 940 - 97) \text{ grams} = 1,311 \text{ grams.} \end{aligned}$$

According to rule (3.4), the uncertainty in this answer is the sum of all four uncertainties,

$$\begin{aligned} \delta M &\approx \delta M_1 + \delta m_1 + \delta M_2 + \delta m_2 = (10 + 1 + 20 + 1) \text{ grams} \\ &= 32 \text{ grams.} \end{aligned}$$

Thus, his final answer (properly rounded) is

$$\text{total mass of liquid} = 1,310 \pm 30 \text{ grams.}$$

Notice how the much smaller uncertainties in the masses of the empty flasks made a negligible contribution to the final uncertainty. This effect is important, and we will discuss it later on. With experience, you can learn to identify in advance those uncertainties that are negligible and can be ignored from the outset. Often, this can greatly simplify the calculation of uncertainties.

PRODUCTS AND QUOTIENTS

Section 2.9 discussed the uncertainty in the product $q = xy$ of two measured quantities. We saw that, provided the fractional uncertainties concerned are small, the fractional uncertainty in $q = xy$ is the sum of the fractional uncertainties in x and y . Rather than review the derivation of this result, I discuss here the similar case of the quotient $q = x/y$. As you will see, the uncertainty in a quotient is given by the same rule as for a product; that is, the fractional uncertainty in $q = x/y$ is equal to the sum of the fractional uncertainties in x and y .

Because uncertainties in products and quotients are best expressed in terms of fractional uncertainties, a shorthand notation for the latter will be helpful. Recall that if we measure some quantity x as

$$(\text{measured value of } x) = x_{\text{best}} \pm \delta x$$

in the usual way, then the fractional uncertainty in x is defined to be

$$(\text{fractional uncertainty in } x) = \frac{\delta x}{|x_{\text{best}}|}.$$

(The absolute value in the denominator ensures that the fractional uncertainty is always positive, even when x_{best} is negative.) Because the symbol $\delta x/|x_{\text{best}}|$ is clumsy to write and read, from now on I will abbreviate it by omitting the subscript “best” and writing

$$(\text{fractional uncertainty in } x) = \frac{\delta x}{|x|}.$$

The result of measuring any quantity x can be expressed in terms of its fractional error $\delta x/|x|$ as

$$(\text{value of } x) = x_{\text{best}}(1 \pm \delta x/|x|).$$

Therefore, the value of $q = x/y$ can be written as

$$(\text{value of } q) = \frac{x_{\text{best}}}{y_{\text{best}}} \frac{1 \pm \delta x/|x|}{1 \pm \delta y/|y|}.$$

Our problem now is to find the extreme probable values of the second factor on the right. This factor is largest, for example, if the numerator has its largest value, $1 + \delta x/|x|$, and the denominator has its *smallest* value, $1 - \delta y/|y|$. Thus, the largest

probable value for $q = x/y$ is

$$\text{(largest value of } q) = \frac{x_{\text{best}}}{y_{\text{best}}} \frac{1 + \delta x/|x|}{1 - \delta y/|y|}. \quad (3.5)$$

The last factor in expression (3.5) has the form $(1 + a)/(1 - b)$, where the numbers a and b are normally small (that is, much less than 1). It can be simplified by two approximations. First, because b is small, the binomial theorem² implies that

$$\frac{1}{(1 - b)} \approx 1 + b. \quad (3.6)$$

Therefore,

$$\begin{aligned} \frac{1 + a}{1 - b} &\approx (1 + a)(1 + b) = 1 + a + b + ab \\ &\approx 1 + a + b, \end{aligned}$$

where, in the second line, we have neglected the product ab of two small quantities. Returning to (3.5) and using these approximations, we find for the largest probable value of $q = x/y$

$$\text{(largest value of } q) = \frac{x_{\text{best}}}{y_{\text{best}}} \left(1 + \frac{\delta x}{|x|} + \frac{\delta y}{|y|} \right).$$

A similar calculation shows that the smallest probable value is given by a similar expression with two minus signs. Combining these two, we find that

$$\text{(value of } q) = \frac{x_{\text{best}}}{y_{\text{best}}} \left(1 \pm \left[\frac{\delta x}{|x|} + \frac{\delta y}{|y|} \right] \right).$$

Comparing this equation with the standard form,

$$\text{(value of } q) = q_{\text{best}} \left(1 \pm \frac{\delta q}{|q|} \right),$$

we see that the best value for q is $q_{\text{best}} = x_{\text{best}}/y_{\text{best}}$, as we would expect, and that the fractional uncertainty is

$$\frac{\delta q}{|q|} \approx \frac{\delta x}{|x|} + \frac{\delta y}{|y|}. \quad (3.7)$$

We conclude that when we divide or multiply two measured quantities x and y , the fractional uncertainty in the answer is the sum of the fractional uncertainties in x and y , as in (3.7). If we now multiply or divide a series of numbers, repeated application of this result leads to the following provisional rule.

²The binomial theorem expresses $1/(1 - b)$ as the infinite series $1 + b + b^2 + \dots$. If b is much less than 1, then $1/(1 - b) \approx 1 + b$ as in (3.6). If you are unfamiliar with the binomial theorem, you can find more details in Problem 3.8.

for... 1.1

Uncertainty in Products and Quotients (Provisional Rule)

If several quantities x, \dots, w are measured with small uncertainties $\delta x, \dots, \delta w$, and the measured values are used to compute

$$q = \frac{x \times \dots \times z}{u \times \dots \times w},$$

then the fractional uncertainty in the computed value of q is the sum,

$$\frac{\delta q}{|q|} \approx \frac{\delta x}{|x|} + \dots + \frac{\delta z}{|z|} + \frac{\delta u}{|u|} + \dots + \frac{\delta w}{|w|}, \quad (3.8)$$

of the fractional uncertainties in x, \dots, w .

Briefly, when quantities are multiplied or divided the *fractional uncertainties add*.

Example: A Problem in Surveying

In surveying, sometimes a value can be found for an inaccessible length l (such as the height of a tall tree) by measuring three other lengths l_1, l_2, l_3 in terms of which

$$l = \frac{l_1 l_2}{l_3}.$$

Suppose we perform such an experiment and obtain the following results (in feet):

$$l_1 = 200 \pm 2, \quad l_2 = 5.5 \pm 0.1, \quad l_3 = 10.0 \pm 0.4.$$

Our best estimate for l is

$$l_{\text{best}} = \frac{200 \times 5.5}{10.0} = 110 \text{ ft.}$$

According to (3.8), the fractional uncertainty in this answer is the sum of the fractional uncertainties in l_1, l_2 , and l_3 , which are 1%, 2%, and 4%, respectively. Thus

$$\begin{aligned} \frac{\delta l}{l} &\approx \frac{\delta l_1}{l_1} + \frac{\delta l_2}{l_2} + \frac{\delta l_3}{l_3} = (1 + 2 + 4)\% \\ &= 7\%, \end{aligned}$$

and our final answer is

$$l = 110 \pm 8 \text{ ft.}$$

Quick Check 3.2. Suppose you measure the three quantities x , y , and z as follows:

$$x = 8.0 \pm 0.2, \quad y = 5.0 \pm 0.1, \quad z = 4.0 \pm 0.1.$$

Express the given uncertainties as percentages, and then calculate $q = xy/z$ with its uncertainty δq [as given by the provisional rule (3.8)].

3.4 Two Important Special Cases

Two important special cases of the rule (3.8) deserve mention. One concerns the product of two numbers, one of which has *no uncertainty*; the other involves a power (such as x^3) of a measured number.

MEASURED QUANTITY TIMES EXACT NUMBER

Suppose we measure a quantity x and then use the measured value to calculate the product $q = Bx$, where the number B has *no uncertainty*. For example, we might measure the diameter of a circle and then calculate its circumference, $c = \pi \times d$; or we might measure the thickness T of 200 identical sheets of paper and then calculate the thickness of a single sheet as $t = (1/200) \times T$. According to the rule (3.8), the fractional uncertainty in $q = Bx$ is the sum of the fractional uncertainties in B and x . Because $\delta B = 0$, this implies that

$$\frac{\delta q}{|q|} = \frac{\delta x}{|x|}.$$

That is, the fractional uncertainty in $q = Bx$ (with B known exactly) is the same as that in x . We can express this result differently if we multiply through by $|q| = |Bx|$ to give $\delta q = |B|\delta x$, and we have the following useful rule:³

Measured Quantity Times Exact Number

If the quantity x is measured with uncertainty δx and is used to compute the product

$$q = Bx,$$

where B has no uncertainty, then the uncertainty in q is just $|B|$ times that in x ,

$$\delta q = |B|\delta x. \tag{3.9}$$

³This rule (3.9) was derived from the rule (3.8), which is provisional and will be replaced by the more complete rules (3.18) and (3.19). Fortunately, the same conclusion (3.9) follows from these improved rules. Thus (3.9) is already in its final form.

This rule is especially useful in measuring something inconveniently small but available many times over, such as the thickness of a sheet of paper or the time for a revolution of a rapidly spinning wheel. For example, if we measure the thickness T of 200 sheets of paper and get the answer

$$(\text{thickness of 200 sheets}) = T = 1.3 \pm 0.1 \text{ inches},$$

it immediately follows that the thickness t of a single sheet is

$$\begin{aligned} (\text{thickness of one sheet}) = t &= \frac{1}{200} \times T \\ &= 0.0065 \pm 0.0005 \text{ inches}. \end{aligned}$$

Notice how this technique (measuring the thickness of several identical sheets and dividing by their number) makes easily possible a measurement that would otherwise require quite sophisticated equipment and that this technique gives a remarkably small uncertainty. Of course, the sheets must be known to be equally thick.

Quick Check 3.3. Suppose you measure the diameter of a circle as

$$d = 5.0 \pm 0.1 \text{ cm}$$

and use this value to calculate the circumference $c = \pi d$. What is your answer, with its uncertainty?

POWERS

The second special case of the rule (3.8) concerns the evaluation of a power of some measured quantity. For example, we might measure the speed v of some object and then, to find its kinetic energy $\frac{1}{2}mv^2$, calculate the square v^2 . Because v^2 is just $v \times v$, it follows from (3.8) that the fractional uncertainty in v^2 is *twice* the fractional uncertainty in v . More generally, from (3.8) the general rule for any power is clearly as follows.

Uncertainty in a Power

If the quantity x is measured with uncertainty δx and the measured value is used to compute the power

$$q = x^n,$$

then the fractional uncertainty in q is n times that in x ,

$$\frac{\delta q}{|q|} = n \frac{\delta x}{|x|}. \quad (3.10)$$

The derivation of this rule required that n be a positive integer. In fact, however, the rule generalizes to include *any* exponent n , as we will see later in Equation (3.26).

Quick Check 3.4. To find the volume of a certain cube, you measure its side as 2.00 ± 0.02 cm. Convert this uncertainty to a percent and then find the volume with its uncertainty.

Example: Measurement of g

Suppose a student measures g , the acceleration of gravity, by measuring the time t for a stone to fall from a height h above the ground. After making several timings, she concludes that

$$t = 1.6 \pm 0.1 \text{ s},$$

and she measures the height h as

$$h = 46.2 \pm 0.3 \text{ ft.}$$

Because h is given by the well-known formula $h = \frac{1}{2}gt^2$, she now calculates g as

$$\begin{aligned} g &= \frac{2h}{t^2} \\ &= \frac{2 \times 46.2 \text{ ft}}{(1.6 \text{ s})^2} = 36.1 \text{ ft/s}^2. \end{aligned}$$

What is the uncertainty in her answer?

The uncertainty in her answer can be found by using the rules just developed. To this end, we need to know the fractional uncertainties in each of the factors in the expression $g = 2h/t^2$ used to calculate g . The factor 2 has no uncertainty. The fractional uncertainties in h and t are

$$\frac{\delta h}{h} = \frac{0.3}{46.2} = 0.7\%$$

and

$$\frac{\delta t}{t} = \frac{0.1}{1.6} = 6.3\%.$$

According to the rule (3.10), the fractional uncertainty of t^2 is twice that of t . Therefore, applying the rule (3.8) for products and quotients to the formula $g = 2h/t^2$, we find the fractional uncertainty

$$\begin{aligned} \frac{\delta g}{g} &= \frac{\delta h}{h} + 2 \frac{\delta t}{t} \\ &= 0.7\% + 2 \times (6.3\%) = 13.3\%, \end{aligned} \tag{3.11}$$

and hence the uncertainty

$$\delta g = (36.1 \text{ ft/s}^2) \times \frac{13.3}{100} = 4.80 \text{ ft/s}^2.$$

Thus, our student's final answer (properly rounded) is

$$g = 36 \pm 5 \text{ ft/s}^2.$$

This example illustrates how simple the estimation of uncertainties can often be. It also illustrates how error analysis tells you not only the size of uncertainties but also how to reduce them. In this example, (3.11) shows that the largest contribution comes from the measurement of the time. If we want a more precise value of g , then the measurement of t must be improved; any attempt to improve the measurement of h will be wasted effort.

Finally, the accepted value of g is 32 ft/s^2 , which lies within our student's margins of error. Thus, she can conclude that her measurement, although not especially accurate, is perfectly consistent with the known value of g .

3.5 Independent Uncertainties in a Sum

The rules presented thus far can be summarized quickly: When measured quantities are added or subtracted, the *uncertainties add*; when measured quantities are multiplied or divided, the *fractional uncertainties add*. In this and the next section, I discuss how, under certain conditions, the uncertainties calculated by using these rules may be unnecessarily large. Specifically, you will see that if the original uncertainties are *independent* and *random*, a more realistic (and smaller) estimate of the final uncertainty is given by similar rules in which the uncertainties (or fractional uncertainties) are *added in quadrature* (a procedure defined shortly).

Let us first consider computing the sum, $q = x + y$, of two numbers x and y that have been measured in the standard form

$$(\text{measured value of } x) = x_{\text{best}} \pm \delta x,$$

with a similar expression for y . The argument used in the last section was as follows: First, the best estimate for $q = x + y$ is obviously $q_{\text{best}} = x_{\text{best}} + y_{\text{best}}$. Second, since the highest probable values for x and y are $x_{\text{best}} + \delta x$ and $y_{\text{best}} + \delta y$, the highest probable value for q is

$$x_{\text{best}} + y_{\text{best}} + \delta x + \delta y. \quad (3.12)$$

Similarly, the lowest probable value of q is

$$x_{\text{best}} + y_{\text{best}} - \delta x - \delta y.$$

Therefore, we concluded, the value of q probably lies between these two numbers, and the uncertainty in q is

$$\delta q \approx \delta x + \delta y.$$

To see why this formula is likely to overestimate δq , let us consider how the actual value of q could equal the highest extreme (3.12). Obviously, this occurs if we have underestimated x by the full amount δx and underestimated y by the full δy , obviously, a fairly unlikely event. If x and y are measured independently and our errors are random in nature, we have a 50% chance that an *underestimate* of x is accompanied by an *overestimate* of y , or *vice versa*. Clearly, then, the probability we will underestimate both x and y by the full amounts δx and δy is fairly small. Therefore, the value $\delta q \approx \delta x + \delta y$ overstates our probable error.

What constitutes a better estimate of δq ? The answer depends on precisely what we mean by uncertainties (that is, what we mean by the statement that q is “probably” somewhere between $q_{\text{best}} - \delta q$ and $q_{\text{best}} + \delta q$). It also depends on the statistical laws governing our errors in measurement. Chapter 5 discusses the normal, or Gauss, distribution, which describes measurements subject to random uncertainties. It shows that if the measurements of x and y are made independently and are both governed by the normal distribution, then the uncertainty in $q = x + y$ is given by

$$\delta q = \sqrt{(\delta x)^2 + (\delta y)^2}. \quad (3.13)$$

When we combine two numbers by squaring them, adding the squares, and taking the square root, as in (3.13), the numbers are said to be *added in quadrature*. Thus, the rule embodied in (3.13) can be stated as follows: If the measurements of x and y are independent and subject only to random uncertainties, then the uncertainty δq in the calculated value of $q = x + y$ is the *sum in quadrature* or *quadratic sum* of the uncertainties δx and δy .

Compare the new expression (3.13) for the uncertainty in $q = x + y$ with our old expression,

$$\delta q \approx \delta x + \delta y. \quad (3.14)$$

First, the new expression (3.13) is always smaller than the old (3.14), as we can see from a simple geometrical argument: For any two positive numbers a and b , the numbers a , b , and $\sqrt{a^2 + b^2}$ are the three sides of a right-angled triangle (Figure 3.2). Because the length of any side of a triangle is always less than the sum of the

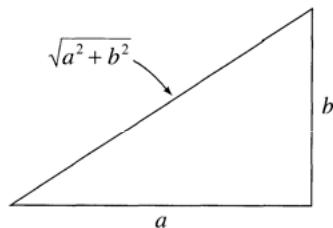


Figure 3.2. Because any side of a triangle is less than the sum of the other two sides, the inequality $\sqrt{a^2 + b^2} < a + b$ is always true.

other two sides, it follows that $\sqrt{a^2 + b^2} < a + b$ and hence that (3.13) is always less than (3.14).

Because expression (3.13) for the uncertainty in $q = x + y$ is always smaller

than (3.14), you should always use (3.13) *when* it is applicable. It is, however, *not* always applicable. Expression (3.13) reflects the possibility that an overestimate of x can be offset by an underestimate of y or vice versa, but there are measurements for which this cancellation is not possible.

Suppose, for example, that $q = x + y$ is the sum of two lengths x and y measured with the same steel tape. Suppose further that the main source of uncertainty is our fear that the tape was designed for use at a temperature different from the present temperature. If we don't know this temperature (and don't have a reliable tape for comparison), we have to recognize that our tape may be longer or shorter than its calibrated length and hence may yield readings under or over the correct length. This uncertainty can be easily allowed for.⁴ The point, however, is that if the tape is too long, then we *underestimate both* x and y ; and if the tape is too short, we *overestimate both* x and y . Thus, there is no possibility for the cancellations that justified using the sum in quadrature to compute the uncertainty in $q = x + y$.

I will prove later (in Chapter 9) that, whether or not our errors are independent and random, the uncertainty in $q = x + y$ is *certainly no larger* than the simple sum $\delta x + \delta y$:

$$\delta q \leq \delta x + \delta y. \quad (3.15)$$

That is, our old expression (3.14) for δq is actually an *upper bound* that holds in all cases. If we have any reason to suspect the errors in x and y are *not* independent and random (as in the example of the steel tape measure), we are not justified in using the quadratic sum (3.13) for δq . On the other hand, the bound (3.15) guarantees that δq is certainly no worse than $\delta x + \delta y$, and our safest course is to use the old rule

$$\delta q \approx \delta x + \delta y.$$

Often, whether uncertainties are added in quadrature or directly makes little difference. For example, suppose that x and y are lengths both measured with uncertainties $\delta x = \delta y = 2$ mm. If we are sure these uncertainties are independent and random, we would estimate the error in $x + y$ to be the sum in quadrature,

$$\sqrt{(\delta x)^2 + (\delta y)^2} = \sqrt{4 + 4} \text{ mm} = 2.8 \text{ mm} \approx 3 \text{ mm},$$

but if we suspect that the uncertainties may not be independent, we would have to use the ordinary sum,

$$\delta x + \delta y \approx (2 + 2) \text{ mm} = 4 \text{ mm}.$$

In many experiments, the estimation of uncertainties is so crude that the difference between these two answers (3 mm and 4 mm) is unimportant. On the other hand, sometimes the sum in quadrature is significantly smaller than the ordinary sum. Also, rather surprisingly, the sum in quadrature is sometimes easier to compute than the ordinary sum. Examples of these effects are given in the next section.

⁴Suppose, for example, that the tape has a coefficient of expansion $\alpha = 10^{-5}$ per degree and that we decide that the difference between its calibration temperature and the present temperature is unlikely to be more than 10 degrees. The tape is then unlikely to be more than 10^{-4} , or 0.01%, away from its correct length, and our uncertainty is therefore 0.01%.

Quick Check 3.5. Suppose you measure the volumes of water in two beakers as

$$V_1 = 130 \pm 6 \text{ ml} \quad \text{and} \quad V_2 = 65 \pm 4 \text{ ml}$$

and then carefully pour the contents of the first into the second. What is your prediction for the total volume $V = V_1 + V_2$ with its uncertainty, δV , assuming the original uncertainties are independent and random? What would you give for δV if you suspected the original uncertainties were not independent?

3.6 More About Independent Uncertainties

In the previous section, I discussed how independent random uncertainties in two quantities x and y propagate to cause an uncertainty in the sum $x + y$. We saw that for this type of uncertainty the two errors should be added in quadrature. We can naturally consider the corresponding problem for differences, products, and quotients. As we will see in Section 5.6, in all cases our previous rules (3.4) and (3.8) are modified only in that the sums of errors (or fractional errors) are replaced by quadratic sums. Further, the old expressions (3.4) and (3.8) will be proven to be upper bounds that always hold whether or not the uncertainties are independent and random. Thus, the final versions of our two main rules are as follows:

Uncertainty in Sums and Differences

Suppose that x, \dots, w are measured with uncertainties $\delta x, \dots, \delta w$ and the measured values used to compute

$$q = x + \dots + z - (u + \dots + w).$$

If the uncertainties in x, \dots, w are known to be *independent and random*, then the uncertainty in q is the quadratic sum

$$\delta q = \sqrt{(\delta x)^2 + \dots + (\delta z)^2 + (\delta u)^2 + \dots + (\delta w)^2} \quad (3.16)$$

of the original uncertainties. In any case, δq is never larger than their ordinary sum,

$$\delta q \leq \delta x + \dots + \delta z + \delta u + \dots + \delta w. \quad (3.17)$$

and

Uncertainties in Products and Quotients

Suppose that x, \dots, w are measured with uncertainties $\delta x, \dots, \delta w$, and the measured values are used to compute

$$q = \frac{x \times \cdots \times z}{u \times \cdots \times w}.$$

If the uncertainties in x, \dots, w are *independent and random*, then the fractional uncertainty in q is the sum in quadrature of the original fractional uncertainties,

$$\frac{\delta q}{|q|} = \sqrt{\left(\frac{\delta x}{x}\right)^2 + \cdots + \left(\frac{\delta z}{z}\right)^2 + \left(\frac{\delta u}{u}\right)^2 + \cdots + \left(\frac{\delta w}{w}\right)^2}. \quad (3.18)$$

In any case, it is never larger than their ordinary sum,

$$\frac{\delta q}{|q|} \leq \frac{\delta x}{|x|} + \cdots + \frac{\delta z}{|z|} + \frac{\delta u}{|u|} + \cdots + \frac{\delta w}{|w|}. \quad (3.19)$$

Notice that I have not yet justified the use of addition in quadrature for independent random uncertainties. I have argued only that when the various uncertainties are independent and random, there is a good chance of partial cancellations of errors and that the resulting uncertainty (or fractional uncertainty) should be smaller than the simple sum of the original uncertainties (or fractional uncertainties); the sum in quadrature does have this property. I give a proper justification of its use in Chapter 5. The bounds (3.17) and (3.19) are proved in Chapter 9.

Example: Straight Addition vs Addition in Quadrature

As discussed, sometimes there is no significant difference between uncertainties computed by addition in quadrature and those computed by straight addition. Often, however, there is a significant difference, and—surprisingly enough—the sum in quadrature is often much simpler to compute. To see how this situation can arise, consider the following example.

Suppose we want to find the efficiency of a D.C. electric motor by using it to lift a mass m through a height h . The work accomplished is mgh , and the electric energy delivered to the motor is VIt , where V is the applied voltage, I the current, and t the time for which the motor runs. The efficiency is then

$$\text{efficiency, } e = \frac{\text{work done by motor}}{\text{energy delivered to motor}} = \frac{mgh}{VIt}.$$

Let us suppose that m , h , V , and I can all be measured with 1% accuracy,

$$(\text{fractional uncertainty for } m, h, V, \text{ and } I) = 1\%,$$

and that the time t has an uncertainty of 5%,

$$(\text{fractional uncertainty for } t) = 5\%.$$

(Of course, g is known with negligible uncertainty.) If we now compute the efficiency e , then according to our old rule (“fractional errors add”), we have an uncertainty

$$\begin{aligned}\frac{\delta e}{e} &\approx \frac{\delta m}{m} + \frac{\delta h}{h} + \frac{\delta V}{V} + \frac{\delta I}{I} + \frac{\delta t}{t} \\ &= (1 + 1 + 1 + 1 + 5)\% = 9\%.\end{aligned}$$

On the other hand, if we are confident that the various uncertainties are independent and random, then we can compute $\delta e/e$ by the quadratic sum to give

$$\begin{aligned}\frac{\delta e}{e} &= \sqrt{\left(\frac{\delta m}{m}\right)^2 + \left(\frac{\delta h}{h}\right)^2 + \left(\frac{\delta V}{V}\right)^2 + \left(\frac{\delta I}{I}\right)^2 + \left(\frac{\delta t}{t}\right)^2} \\ &= \sqrt{(1\%)^2 + (1\%)^2 + (1\%)^2 + (1\%)^2 + (5\%)^2} \\ &= \sqrt{29\%} \approx 5\%.\end{aligned}$$

Clearly, the quadratic sum leads to a significantly smaller estimate for δe . Furthermore, to one significant figure, the uncertainties in m , h , V , and I *make no contribution at all* to the uncertainty in e computed in this way; that is, to one significant figure, we have found (in this example)

$$\frac{\delta e}{e} = \frac{\delta t}{t}.$$

This striking simplification is easily understood. When numbers are added in quadrature, they are squared first and then summed. The process of squaring greatly exaggerates the importance of the larger numbers. Thus, if one number is 5 times any of the others (as in our example), its square is 25 times that of the others, and we can usually neglect the others entirely.

This example illustrates how combining errors in quadrature is usually better and often easier than computing them by straight addition. The example also illustrates the type of problem in which the errors *are* independent and for which addition in quadrature is justified. (For the moment I take for granted that the errors are random and will discuss this more difficult point in Chapter 4.) The five quantities measured (m , h , V , I , and t) are physically distinct quantities with different units and are measured by entirely different processes. For the sources of error in any quantity to be correlated with those in any other is almost inconceivable. Therefore, the errors can reasonably be treated as independent and combined in quadrature.

Quick Check 3.6. Suppose you measure three numbers as follows:

$$x = 200 \pm 2, \quad y = 50 \pm 2, \quad z = 20 \pm 1,$$

where the three uncertainties are independent and random. What would you give for the values of $q = x + y - z$ and $r = xy/z$ with their uncertainties?

3.7 Arbitrary Functions of One Variable

You have now seen how uncertainties, both independent and otherwise, propagate through sums, differences, products, and quotients. However, many calculations require more complicated operations, such as computation of a sine, cosine, or square root, and you will need to know how uncertainties propagate in these cases.

As an example, imagine finding the refractive index n of glass by measuring the critical angle θ . We know from elementary optics that $n = 1/\sin \theta$. Therefore, if we can measure the angle θ , we can easily calculate the refractive index n , but we must then decide what uncertainty δn in $n = 1/\sin \theta$ results from the uncertainty $\delta\theta$ in our measurement of θ .

More generally, suppose we have measured a quantity x in the standard form $x_{\text{best}} \pm \delta x$ and want to calculate some known function $q(x)$, such as $q(x) = 1/\sin x$ or $q(x) = \sqrt{x}$. A simple way to think about this calculation is to draw a graph of $q(x)$ as in Figure 3.3. The best estimate for $q(x)$ is, of course, $q_{\text{best}} = q(x_{\text{best}})$, and the values x_{best} and q_{best} are shown connected by the heavy lines in Figure 3.3.

To decide on the uncertainty δq , we employ the usual argument. The largest probable value of x is $x_{\text{best}} + \delta x$; using the graph, we can immediately find the largest probable value of q , which is shown as q_{max} . Similarly, we can draw in the smallest probable value, q_{min} , as shown. If the uncertainty δx is small (as we always suppose it is), then the section of graph involved in this construction is approximately straight, and q_{max} and q_{min} are easily seen to be equally spaced on either side of q_{best} . The uncertainty δq can then be taken from the graph as either of the lengths shown, and we have found the value of q in the standard form $q_{\text{best}} \pm \delta q$.

Occasionally, uncertainties are calculated from a graph as just described. (See Problems 3.26 and 3.30 for examples.) Usually, however, the function $q(x)$ is known

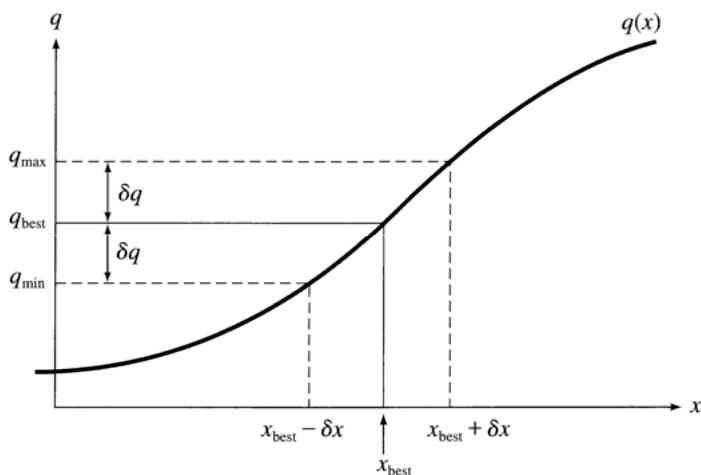


Figure 3.3. Graph of $q(x)$ vs x . If x is measured as $x_{\text{best}} \pm \delta x$, then the best estimate for $q(x)$ is $q_{\text{best}} = q(x_{\text{best}})$. The largest and smallest probable values of $q(x)$ correspond to the values $x_{\text{best}} \pm \delta x$ of x .

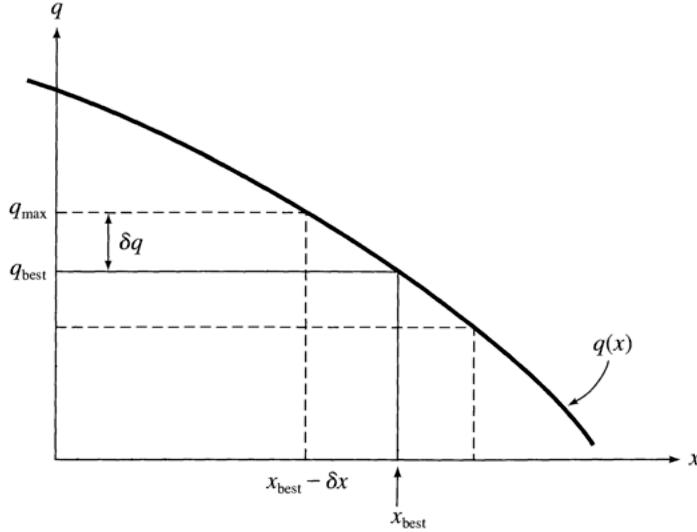


Figure 3.4. If the slope of $q(x)$ is negative, the maximum probable value of q corresponds to the minimum value of x , and vice versa.

explicitly— $q(x) = \sin x$ or $q(x) = \sqrt{x}$, for example—and the uncertainty δq can be calculated analytically. From Figure 3.3, we see that

$$\delta q = q(x_{\text{best}} + \delta x) - q(x_{\text{best}}). \quad (3.20)$$

Now, a fundamental approximation of calculus asserts that, for any function $q(x)$ and any sufficiently small increment u ,

$$q(x + u) - q(x) = \frac{dq}{dx} u.$$

Thus, provided the uncertainty δx is small (as we always assume it is), we can rewrite the difference in (3.20) to give

$$\delta q = \frac{dq}{dx} \delta x. \quad (3.21)$$

Thus, to find the uncertainty δq , we just calculate the derivative dq/dx and multiply by the uncertainty δx .

The rule (3.21) is not quite in its final form. It was derived for a function, like that of Figure 3.3, whose slope is positive. Figure 3.4 shows a function with negative slope. Here, the maximum probable value q_{max} obviously corresponds to the minimum value of x , so that

$$\delta q = -\frac{dq}{dx} \delta x. \quad (3.22)$$

Because dq/dx is negative, we can write $-dq/dx$ as $|dq/dx|$, and we have the following general rule.

Uncertainty in Any Function of One Variable

If x is measured with uncertainty δx and is used to calculate the function $q(x)$, then the uncertainty δq is

$$\delta q = \left| \frac{dq}{dx} \right| \delta x. \quad (3.23)$$

This rule usually allows us to find δq quickly and easily. Occasionally, if $q(x)$ is very complicated, evaluating its derivative may be a nuisance, and going back to (3.20) is sometimes easier, as we discuss in Problem 3.32. Particularly if you have programmed your calculator or computer to find $q(x)$, then finding $q(x_{\text{best}} + \delta x)$ and $q(x_{\text{best}})$ and their difference may be easier than differentiating $q(x)$ explicitly.

Example: Uncertainty in a Cosine

As a simple application of the rule (3.23), suppose we have measured an angle θ as

$$\theta = 20 \pm 3^\circ$$

and that we wish to find $\cos \theta$. Our best estimate of $\cos \theta$ is, of course, $\cos 20^\circ = 0.94$, and according to (3.23), the uncertainty is

$$\begin{aligned} \delta(\cos \theta) &= \left| \frac{d \cos \theta}{d \theta} \right| \delta \theta \\ &= |\sin \theta| \delta \theta \text{ (in rad).} \end{aligned} \quad (3.24)$$

We have indicated that $\delta \theta$ must be expressed in radians, because the derivative of $\cos \theta$ is $-\sin \theta$ only if θ is expressed in radians. Therefore, we rewrite $\delta \theta = 3^\circ$ as $\delta \theta = 0.05$ rad; then (3.24) gives

$$\begin{aligned} \delta(\cos \theta) &= (\sin 20^\circ) \times 0.05 \\ &= 0.34 \times 0.05 = 0.02. \end{aligned}$$

Thus, our final answer is

$$\cos \theta = 0.94 \pm 0.02.$$

Quick Check 3.7. Suppose you measure x as 3.0 ± 0.1 and then calculate $q = e^x$. What is your answer, with its uncertainty? (Remember that the derivative of e^x is e^x .)

As another example of the rule (3.23), we can rederive and generalize a result found in Section 3.4. Suppose we measure the quantity x and then calculate the

power $q(x) = x^n$, where n is any known, fixed number, positive or negative. According to (3.23), the resulting uncertainty in q is

$$\delta q = \left| \frac{dq}{dx} \right| \delta x = |nx^{n-1}| \delta x.$$

If we divide both sides of this equation by $|q| = |x^n|$, we find that

$$\frac{\delta q}{|q|} = |n| \frac{\delta x}{|x|}; \quad (3.25)$$

that is, the fractional uncertainty in $q = x^n$ is $|n|$ times that in x . This result (3.25) is just the rule (3.10) found earlier, except that the result here is more general, because n can now be any number. For example, if $n = 1/2$, then $q = \sqrt{x}$, and

$$\frac{\delta q}{|q|} = \frac{1}{2} \frac{\delta x}{|x|};$$

that is, the fractional uncertainty in \sqrt{x} is *half* that in x itself. Similarly, the fractional uncertainty in $1/x = x^{-1}$ is the same as that in x itself.

The result (3.25) is just a special case of the rule (3.23). It is sufficiently important, however, to deserve separate statement as the following general rule.

Uncertainty in a Power

If x is measured with uncertainty δx and is used to calculate the power $q = x^n$ (where n is a fixed, known number), then the fractional uncertainty in q is $|n|$ times that in x ,

$$\frac{\delta q}{|q|} = |n| \frac{\delta x}{|x|}. \quad (3.26)$$

Quick Check 3.8. If you measure x as 100 ± 6 , what should you report for \sqrt{x} , with its uncertainty?

3.8 Propagation Step by Step

We now have enough tools to handle almost any problem in the propagation of errors. Any calculation can be broken down into a sequence of steps, each involving just one of the following types of operation: (1) sums and differences; (2) products and quotients; and (3) computation of a function of one variable, such as x^n , $\sin x$,

e^x , or $\ln x$. For example, we could calculate

$$q = x(y - z \sin u) \quad (3.27)$$

from the measured quantities x , y , z , and u in the following steps: Compute the *function* $\sin u$, then the *product* of z and $\sin u$, next the *difference* of y and $z \sin u$, and finally the *product* of x and $(y - z \sin u)$.

We know how uncertainties propagate through each of these separate operations. Thus, provided the various quantities involved are independent, we can calculate the uncertainty in the final answer by proceeding in steps from the uncertainties in the original measurement. For example, if the quantities x , y , z , and u in (3.27) have been measured with corresponding uncertainties $\delta x, \dots, \delta u$, we could calculate the uncertainty in q as follows. First, find the uncertainty in the function $\sin u$; knowing this, find the uncertainty in the product $z \sin u$, and then that in the difference $y - z \sin u$; finally, find the uncertainty in the complete product (3.27).

Quick Check 3.9. Suppose you measure three numbers as follows:

$$x = 200 \pm 2, \quad y = 50 \pm 2, \quad z = 40 \pm 2,$$

where the three uncertainties are independent and random. Use step-by-step propagation to find the quantity $q = x/(y - z)$ with its uncertainty. [First find the uncertainty in the difference $y - z$ and then the quotient $x/(y - z)$.]

Before I discuss some examples of this step-by-step calculation of errors, let me emphasize three general points. First, because uncertainties in sums or differences involve absolute uncertainties (such as δx) whereas those in products or quotients involve fractional uncertainties (such as $\delta x/|x|$), the calculations will require some facility in passing from absolute to fractional uncertainties and vice versa, as demonstrated below.

Second, an important simplifying feature of all these calculations is that (as repeatedly emphasized) uncertainties are seldom needed to more than one significant figure. Hence, much of the calculation can be done rapidly in your head, and many smaller uncertainties can be completely neglected. In a typical experiment involving several trials, you may need to do a careful calculation on paper of all error propagations for the first trial. After that, you will often find that all trials are sufficiently similar that no further calculation is needed or, at worst, that for subsequent trials the calculations of the first trial can be modified in your head.

Finally, you need to be aware that you will sometimes encounter functions $q(x)$ whose uncertainty cannot be found reliably by the stepwise method advocated here. These functions always involve at least one variable that appears more than once. Suppose, for example, that in place of the function (3.27), we had to evaluate

$$q = y - x \sin y.$$

This function is the difference of two terms, y and $x \sin y$, but these two terms are definitely *not* independent because both depend on y . Thus, to estimate the uncertainty, we would have to treat the terms as dependent (that is, add their uncertainties directly, not in quadrature). Under some circumstances, this treatment may seriously overestimate the true uncertainty. Faced with a function like this, we must recognize that a stepwise calculation may give an uncertainty that is unnecessarily big, and the only satisfactory procedure is then to use the general formula to be developed in Section 3.11.

3.9 Examples

In this and the next section, I give three examples of the type of calculation encountered in introductory laboratories. None of these examples is especially complicated; in fact, few real problems are much more complicated than the ones described here.

Example: Measurement of g with a Simple Pendulum

As a first example, suppose that we measure g , the acceleration of gravity, using a simple pendulum. The period of such a pendulum is well known to be $T = 2\pi\sqrt{l/g}$, where l is the length of the pendulum. Thus, if l and T are measured, we can find g as

$$g = 4\pi^2 l/T^2. \quad (3.28)$$

This result gives g as the product or quotient of three factors, $4\pi^2$, l , and T^2 . If the various uncertainties are independent and random, the fractional uncertainty in our answer is just the quadratic sum of the fractional uncertainties in these factors. The factor $4\pi^2$ has no uncertainty, and the fractional uncertainty in T^2 is twice that in T :

$$\frac{\delta(T^2)}{T^2} = 2 \frac{\delta T}{T}.$$

Thus, the fractional uncertainty in our answer for g will be

$$\frac{\delta g}{g} = \sqrt{\left(\frac{\delta l}{l}\right)^2 + \left(2 \frac{\delta T}{T}\right)^2}. \quad (3.29)$$

Suppose we measure the period T for one value of the length l and get the results⁵

$$l = 92.95 \pm 0.1 \text{ cm},$$

$$T = 1.936 \pm 0.004 \text{ s}.$$

⁵Although at first sight an uncertainty $\delta T = 0.004$ s may seem unrealistically small, you can easily achieve it by timing several oscillations. If you can measure with an accuracy of 0.1 s, as is certainly possible with a stopwatch, then by timing 25 oscillations you will find T within 0.004 s.

Our best estimate for g is easily found from (3.28) as

$$g_{\text{best}} = \frac{4\pi^2 \times (92.95 \text{ cm})}{(1.936 \text{ s})^2} = 979 \text{ cm/s}^2.$$

To find our uncertainty in g using (3.29), we need the fractional uncertainties in l and T . These are easily calculated (in the head) as

$$\frac{\delta l}{l} = 0.1\% \quad \text{and} \quad \frac{\delta T}{T} = 0.2\%.$$

Substituting into (3.29), we find

$$\frac{\delta g}{g} = \sqrt{(0.1)^2 + (2 \times 0.2)^2}\% = 0.4\%;$$

from which

$$\delta g = 0.004 \times 979 \text{ cm/s}^2 = 4 \text{ cm/s}^2.$$

Thus, based on these measurements, our final answer is

$$g = 979 \pm 4 \text{ cm/s}^2.$$

Having found the measured value of g and its uncertainty, we would naturally compare these values with the accepted value of g . If the latter has its usual value of 981 cm/s^2 , the present value is entirely satisfactory.

If this experiment is repeated (as most such experiments should be) with different values of the parameters, the uncertainty calculations usually do not need to be repeated in complete detail. We can often easily convince ourselves that all uncertainties (in the answers for g) are close enough that no further calculations are needed; sometimes the uncertainty in a few representative values of g can be calculated and the remainder estimated by inspection. In any case, the best procedure is almost always to record the various values of l , T , and g and the corresponding uncertainties in a single table. (See Problem 3.40.)

Example: Refractive Index Using Snell's Law

If a ray of light passes from air into glass, the angles of incidence i and refraction r are defined as in Figure 3.5 and are related by Snell's law, $\sin i = n \sin r$, where n is the refractive index of the glass. Thus, if you measure the angles i and r , you

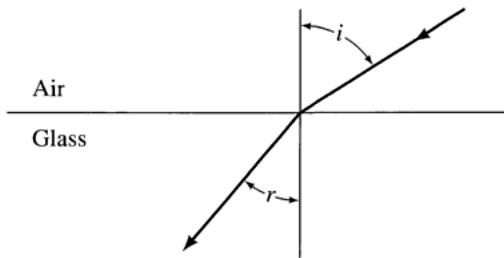


Figure 3.5. The angles of incidence i and refraction r when a ray of light passes from air into glass.

can calculate the refractive index n as

$$n = \frac{\sin i}{\sin r}. \quad (3.30)$$

The uncertainty in this answer is easily calculated. Because n is the quotient of $\sin i$ and $\sin r$, the fractional uncertainty in n is the quadratic sum of those in $\sin i$ and $\sin r$:

$$\frac{\delta n}{n} = \sqrt{\left(\frac{\delta \sin i}{\sin i}\right)^2 + \left(\frac{\delta \sin r}{\sin r}\right)^2}. \quad (3.31)$$

To find the fractional uncertainty in the sine of any angle θ , we note that

$$\begin{aligned} \delta \sin \theta &= \left| \frac{d \sin \theta}{d \theta} \right| \delta \theta \\ &= |\cos \theta| \delta \theta \text{ (in rad).} \end{aligned}$$

Thus, the fractional uncertainty is

$$\frac{\delta \sin \theta}{|\sin \theta|} = |\cot \theta| \delta \theta \text{ (in rad).} \quad (3.32)$$

Suppose we now measure the angle r for a couple of values of i and get the results shown in the first two columns of Table 3.1 (with all measurements judged to be uncertain by $\pm 1^\circ$, or 0.02 rad). The calculation of $n = \sin i / \sin r$ is easily carried out as shown in the next three columns of Table 3.1. The uncertainty in n can then be found as in the last three columns; the fractional uncertainties in $\sin i$ and $\sin r$ are calculated using (3.32), and finally the fractional uncertainty in n is found using (3.31).

Table 3.1. Finding the refractive index.

i (deg) all ± 1	r (deg) all ± 1	$\sin i$	$\sin r$	n	$\frac{\delta \sin i}{ \sin i }$	$\frac{\delta \sin r}{ \sin r }$	$\frac{\delta n}{n}$
20	13	0.342	0.225	1.52	5%	8%	9%
40	23.5	0.643	0.399	1.61	2%	4%	5%

Before making a series of measurements like the two shown in Table 3.1, you should think carefully how best to record the data and calculations. A tidy display like that in Table 3.1 makes the recording of data easier and reduces the danger of mistakes in calculation. It is also easier for the reader to follow and check.

If you repeat an experiment like this one several times, the error calculations can become tedious if you do them for each repetition. If you have a programmable calculator, you may decide to write a program to do the repetitive calculations automatically. You should recognize, however, that you almost never need to do the error calculations for all the repetitions; if you find the uncertainties in n corresponding to the smallest and largest values of i (and possibly a few intermediate values), then these uncertainties suffice for most purposes.

3.10 A More Complicated Example

The two examples just given are typical of many experiments in the introductory physics laboratory. A few experiments require more complicated calculations, however. As an example of such an experiment, I discuss here the measurement of the acceleration of a cart rolling down a slope.⁶

Example: Acceleration of a Cart Down a Slope

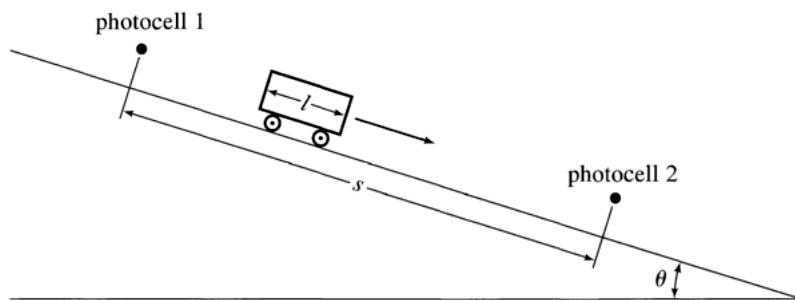


Figure 3.6. A cart rolls down an incline of slope θ . Each photocell is connected to a timer to measure the time for the cart to pass it.

Let us consider a cart rolling down an incline of slope θ as in Figure 3.6. The expected acceleration is $g \sin \theta$ and, if we measure θ , we can easily calculate the expected acceleration and its uncertainty (Problem 3.42). We can measure the actual acceleration a by timing the cart past two photocells as shown, each connected to a timer. If the cart has length l and takes time t_1 to pass the first photocell, its speed there is $v_1 = l/t_1$. In the same way, $v_2 = l/t_2$. (Strictly speaking, these speeds are the cart's *average* speeds while passing the two photocells. However, provided l is small, the difference between the average and instantaneous speeds is unimportant.) If the distance between the photocells is s , then the well-known formula $v_2^2 = v_1^2 + 2as$ implies that

$$\begin{aligned} a &= \frac{v_2^2 - v_1^2}{2s} \\ &= \left(\frac{l^2}{2s} \right) \left(\frac{1}{t_2^2} - \frac{1}{t_1^2} \right). \end{aligned} \quad (3.33)$$

Using this formula and the measured values of l , s , t_1 , and t_2 , we can easily find the observed acceleration and its uncertainty.

⁶If you wish, you could omit this section without loss of continuity or return to study it in connection with Problem 3.42.

One set of data for this experiment, including uncertainties, was as follows (the numbers in parentheses are the corresponding percentage uncertainties, as you can easily check):

$$\begin{aligned} l &= 5.00 \pm 0.05 \text{ cm (1\%)} \\ s &= 100.0 \pm 0.2 \text{ cm (0.2\%)} \\ t_1 &= 0.054 \pm 0.001 \text{ s (2\%)} \\ t_2 &= 0.031 \pm 0.001 \text{ s (3\%).} \end{aligned} \quad (3.34)$$

From these values, we can immediately calculate the first factor in (3.33) as $l^2/2s = 0.125 \text{ cm}$. Because the fractional uncertainties in l and s are 1% and 0.2%, that in $l^2/2s$ is

$$\begin{aligned} (\text{fractional uncertainty in } l^2/2s) &= \sqrt{\left(2 \frac{\delta l}{l}\right)^2 + \left(\frac{\delta s}{s}\right)^2} \\ &= \sqrt{(2 \times 1\%)^2 + (0.2\%)^2} = 2\%. \end{aligned}$$

(Note how the uncertainty in s makes no appreciable contribution and could have been ignored.) Therefore,

$$l^2/2s = 0.125 \text{ cm} \pm 2\%. \quad (3.35)$$

To calculate the second factor in (3.33) and its uncertainty, we proceed in steps. Because the fractional uncertainty in t_1 is 2%, that in $1/t_1^2$ is 4%. Thus, since $t_1 = 0.054 \text{ s}$,

$$1/t_1^2 = 343 \pm 14 \text{ s}^{-2}.$$

In the same way, the fractional uncertainty in $1/t_2^2$ is 6% and

$$1/t_2^2 = 1041 \pm 62 \text{ s}^{-2}.$$

Subtracting these (and combining the errors in quadrature), we find

$$\frac{1}{t_2^2} - \frac{1}{t_1^2} = 698 \pm 64 \text{ s}^{-2} \quad (\text{or } 9\%). \quad (3.36)$$

Finally, according to (3.33), the required acceleration is the product of (3.35) and (3.36). Multiplying these equations together (and combining the fractional uncertainties in quadrature), we obtain

$$\begin{aligned} a &= (0.125 \text{ cm} \pm 2\%) \times (698 \text{ s}^{-2} \pm 9\%) \\ &= 87.3 \text{ cm/s}^2 \pm 9\% \end{aligned}$$

or

$$a = 87 \pm 8 \text{ cm/s}^2. \quad (3.37)$$

This answer could now be compared with the expected acceleration $g \sin \theta$, if the latter had been calculated.

When the calculations leading to (3.37) are studied carefully, several interesting features emerge. First, the 2% uncertainty in the factor $l^2/2s$ is completely swamped

by the 9% uncertainty in $(1/t_2^2) - (1/t_1^2)$. If further calculations are needed for subsequent trials, the uncertainties in t and s can therefore be ignored (so long as a quick check shows they are still just as unimportant).

Another important feature of our calculation is the way in which the 2% and 3% uncertainties in t_1 and t_2 grow when we evaluate $1/t_1^2$, $1/t_2^2$, and the difference $(1/t_2^2) - (1/t_1^2)$, so that the final uncertainty is 9%. This growth results partly from taking squares and partly from taking the difference of large numbers. We could imagine extending the experiment to check the constancy of a by giving the cart an initial push, so that the speeds v_1 and v_2 are both larger. If we did, the times t_1 and t_2 would get smaller, and the effects just described would get worse (see Problem 3.42).

3.11 General Formula for Error Propagation⁷

So far, we have established three main rules for the propagation of errors: that for sums and differences, that for products and quotients, and that for arbitrary functions of one variable. In the past three sections, we have seen how the computation of a complicated function can often be broken into steps and the uncertainty in the function computed stepwise using our three simple rules.

In this final section, I give a single general formula from which all three of these rules can be derived and with which any problem in error propagation can be solved. Although this formula is often rather cumbersome to use, it is useful theoretically. Furthermore, there are some problems in which, instead of calculating the uncertainty in steps as in the past three sections, you will do better to calculate it in one step by means of the general formula.

To illustrate the kind of problem for which the one-step calculation is preferable, suppose that we measure three quantities x , y , and z and have to compute a function such as

$$q = \frac{x + y}{x + z} \quad (3.38)$$

in which a variable appears more than once (x in this case). If we were to calculate the uncertainty δq in steps, then we would first compute the uncertainties in the two sums $x + y$ and $x + z$, and then that in their quotient. Proceeding in this way, we would completely miss the possibility that errors in the numerator due to errors in x may, to some extent, cancel errors in the denominator due to errors in x . To understand how this cancellation can happen, suppose that x , y , and z are all positive numbers, and consider what happens if our measurement of x is subject to error. If we overestimate x , we overestimate both $x + y$ and $x + z$, and (to a large extent) these overestimates cancel one another when we calculate $(x + y)/(x + z)$. Similarly, an underestimate of x leads to underestimates of both $x + y$ and $x + z$, which again cancel when we form the quotient. In either case, an error in x is substantially

⁷You can postpone reading this section without a serious loss of continuity. The material covered here is not used again until Section 5.6.

canceled out of the quotient $(x + y)/(x + z)$, and our stepwise calculation completely misses these cancellations.

Whenever a function involves the same quantity more than once, as in (3.38), some errors may cancel themselves (an effect sometimes called *compensating errors*). If this cancellation is possible, then a stepwise calculation of the uncertainty may overestimate the final uncertainty. The only way to avoid this overestimation is to calculate the uncertainty in one step by using the method I will now develop.⁸

Let us suppose at first that we measure two quantities x and y and then calculate some function $q = q(x, y)$. This function could be as simple as $q = x + y$ or something more complicated such as $q = (x^3 + y) \sin(xy)$. For a function $q(x)$ of a *single* variable, we argued that if the best estimate for x is the number x_{best} , then the best estimate for $q(x)$ is $q(x_{\text{best}})$. Next, we argued that the extreme (that is, largest and smallest) probable values of x are $x_{\text{best}} \pm \delta x$ and that the corresponding extreme values of q are therefore

$$q(x_{\text{best}} \pm \delta x). \quad (3.39)$$

Finally, we used the approximation

$$q(x + u) \approx q(x) + \frac{dq}{dx} u \quad (3.40)$$

(for any small increment u) to rewrite the extreme probable values (3.39) as

$$q(x_{\text{best}}) \pm \left| \frac{dq}{dx} \right| \delta x, \quad (3.41)$$

where the absolute value is to allow for the possibility that dq/dx may be negative. The result (3.41) means that $\delta q \approx |dq/dx|\delta x$.

When q is a function of two variables, $q(x, y)$, the argument is similar. If x_{best} and y_{best} are the best estimates for x and y , we expect the best estimate for q to be

$$q_{\text{best}} = q(x_{\text{best}}, y_{\text{best}})$$

in the usual way. To estimate the uncertainty in this result, we need to generalize the approximation (3.40) for a function of two variables. The required generalization is

$$q(x + u, y + v) \approx q(x, y) + \frac{\partial q}{\partial x} u + \frac{\partial q}{\partial y} v, \quad (3.42)$$

where u and v are any small increments in x and y , and $\partial q/\partial x$ and $\partial q/\partial y$ are the so-called *partial derivatives* of q with respect to x and y . That is, $\partial q/\partial x$ is the result of differentiating q with respect to x while treating y as fixed, and vice versa for $\partial q/\partial y$. [For further discussion of partial derivatives and the approximation (3.42), see Problems 3.43 and 3.44.]

The extreme probable values for x and y are $x_{\text{best}} \pm \delta x$ and $y_{\text{best}} \pm \delta y$. If we insert these values into (3.42) and recall that $\partial q/\partial x$ and $\partial q/\partial y$ may be positive or

⁸Sometimes a function that involves a variable more than once can be rewritten in a different form that does not. For example, $q = xy - xz$ can be rewritten as $q = x(y - z)$. In the second form, the uncertainty δq can be calculated in steps without any danger of overestimation.

negative, we find, for the extreme values of q ,

$$q(x_{\text{best}}, y_{\text{best}}) \pm \left(\left| \frac{\partial q}{\partial x} \right| \delta x + \left| \frac{\partial q}{\partial y} \right| \delta y \right).$$

This means that the uncertainty in $q(x, y)$ is

$$\delta q \approx \left| \frac{\partial q}{\partial x} \right| \delta x + \left| \frac{\partial q}{\partial y} \right| \delta y. \quad (3.43)$$

Before I discuss various generalizations of this new rule, let us apply it to rederive some familiar cases. Suppose, for instance, that

$$q(x, y) = x + y; \quad (3.44)$$

that is, q is just the sum of x and y . The partial derivatives are both one,

$$\frac{\partial q}{\partial x} = \frac{\partial q}{\partial y} = 1, \quad (3.45)$$

and so, according to (3.43),

$$\delta q \approx \delta x + \delta y. \quad (3.46)$$

This is just our original provisional rule that the uncertainty in $x + y$ is the sum of the uncertainties in x and y .

In much the same way, if q is the product $q = xy$, you can check that (3.43) implies the familiar rule that the fractional uncertainty in q is the sum of the fractional uncertainties in x and y (see Problem 3.45).

The rule (3.43) can be generalized in various ways. You will not be surprised to learn that when the uncertainties δx and δy are independent and random, the sum (3.43) can be replaced by a sum in quadrature. If the function q depends on more than two variables, then we simply add an extra term for each extra variable. In this way, we arrive at the following general rule (whose full justification will appear in Chapters 5 and 9).

Uncertainty in a Function of Several Variables

Suppose that x, \dots, z are measured with uncertainties $\delta x, \dots, \delta z$ and the measured values are used to compute the function $q(x, \dots, z)$. If the uncertainties in x, \dots, z are independent and random, then the uncertainty in q is

$$\delta q = \sqrt{\left(\frac{\partial q}{\partial x} \delta x \right)^2 + \dots + \left(\frac{\partial q}{\partial z} \delta z \right)^2}. \quad (3.47)$$

In any case, it is never larger than the ordinary sum

$$\delta q \leq \left| \frac{\partial q}{\partial x} \right| \delta x + \dots + \left| \frac{\partial q}{\partial z} \right| \delta z. \quad (3.48)$$

Although the formulas (3.47) and (3.48) look fairly complicated, they are easy to understand if you think about them one term at a time. For example, suppose for a moment that among all the measured quantities, x, y, \dots, z , only x is subject to any uncertainty. (That is, $\delta y = \dots = \delta z = 0$.) Then (3.47) contains only one term and we would find

$$\delta q = \left| \frac{\partial q}{\partial x} \right| \delta x \quad (\text{if } \delta y = \dots = \delta z = 0). \quad (3.49)$$

In other words, the term $|\partial q/\partial x|\delta x$ by itself is the uncertainty, or partial uncertainty, in q caused by the uncertainty in x alone. In the same way, $|\partial q/\partial y|\delta y$ is the partial uncertainty in q due to δy alone, and so on. Referring back to (3.47), we see that the total uncertainty in q is the quadratic sum of the partial uncertainties due to each of the separate uncertainties $\delta x, \delta y, \dots, \delta z$ (provided the latter are independent). This is a good way to think about the result (3.47), and it suggests the simplest way to use (3.47) to calculate the total uncertainty in q : First, calculate the partial uncertainties in q due to $\delta x, \delta y, \dots, \delta z$ separately, using (3.49) and its analogs for y, \dots, z ; then simply combine these separate uncertainties in quadrature to give the total uncertainty as in (3.47).

In the same way, whether or not the uncertainties $\delta x, \delta y, \dots, \delta z$ are independent, the rule (3.48) says that the total uncertainty in q never exceeds the simple sum of the partial uncertainties due to each of $\delta x, \delta y, \dots, \delta z$ separately.

Example: Using the General Formula (3.47)

To determine the quantity

$$q = x^2y - xy^2,$$

a scientist measures x and y as follows:

$$x = 3.0 \pm 0.1 \quad \text{and} \quad y = 2.0 \pm 0.1.$$

What is his answer for q and its uncertainty, as given by (3.47)?

His best estimate for q is easily seen to be $q_{\text{best}} = 6.0$. To find δq , we follow the steps just outlined. The uncertainty in q due to δx alone, which we denote by δq_x , is given by (3.49) as

$$\begin{aligned} \delta q_x &= (\text{error in } q \text{ due to } \delta x \text{ alone}) \\ &= \left| \frac{\partial q}{\partial x} \right| \delta x \\ &= |2xy - y^2| \delta x = |12 - 4| \times 0.1 = 0.8. \end{aligned} \quad (3.50)$$

Similarly, the uncertainty in q due to δy is

$$\begin{aligned} \delta q_y &= (\text{error in } q \text{ due to } \delta y \text{ alone}) \\ &= \left| \frac{\partial q}{\partial y} \right| \delta y \\ &= |x^2 - 2xy| \delta y = |9 - 12| \times 0.1 = 0.3. \end{aligned} \quad (3.51)$$

Finally, according to (3.47), the total uncertainty in q is the quadratic sum of these two partial uncertainties:

$$\begin{aligned}\delta q &= \sqrt{(\delta q_x)^2 + (\delta q_y)^2} \\ &= \sqrt{(0.8)^2 + (0.3)^2} = 0.9.\end{aligned}\quad (3.52)$$

Thus, the final answer for q is

$$q = 6.0 \pm 0.9.$$

The use of (3.47) or (3.48) to calculate uncertainties is reasonably straightforward if you follow the procedure used in this example; that is, first calculate each separate contribution to δq and only then combine them to give the total uncertainty. This procedure breaks the problem into calculations small enough that you have a good chance of getting them right. It has the further advantage that it lets you see which of the measurements x, y, \dots, z are the main contributors to the final uncertainty. (For instance, in the example above, the contribution $\delta q_y = 0.3$ was so small compared with $\delta q_x = 0.8$ that the former could almost be ignored.)

Generally speaking, when the stepwise propagation described in Sections 3.8 to 3.10 is possible, it is usually simpler than the general rules (3.47) or (3.48) discussed here. Nevertheless, you must recognize that if the function $q(x, \dots, z)$ involves any variable more than once, there may be compensating errors; if so, a stepwise calculation may overestimate the final uncertainty, and calculating δq in one step using (3.47) or (3.48) is better.

Principal Definitions and Equations of Chapter 3

THE SQUARE-ROOT RULE FOR A COUNTING EXPERIMENT

If we observe the occurrences of an event that happens at random but with a definite average rate and we count ν occurrences in a time T , our estimate for the true average number is

$$(\text{average number of events in time } T) = \nu \pm \sqrt{\nu}. \quad [\text{See (3.2)}]$$

RULES FOR ERROR PROPAGATION

The rules of error propagation refer to a situation in which we have found various quantities, x, \dots, w with uncertainties $\delta x, \dots, \delta w$ and then use these values to calculate a quantity q . The uncertainties in x, \dots, w "propagate" through the calculation to cause an uncertainty in q as follows:

Sums and Differences: If

$$q = x + \cdots + z - (u + \cdots + w),$$

then

$$\delta q = \sqrt{(\delta x)^2 + \cdots + (\delta z)^2 + (\delta u)^2 + \cdots + (\delta w)^2}$$

(provided all errors are independent and random)

and

$$\delta q \leq \delta x + \cdots + \delta z + \delta u + \cdots + \delta w$$

(always).

[See (3.16) & (3.17)]

Products and Quotients: If

$$q = \frac{x \times \cdots \times z}{u \times \cdots \times w},$$

then

$$\frac{\delta q}{|q|} = \sqrt{\left(\frac{\delta x}{x}\right)^2 + \cdots + \left(\frac{\delta z}{z}\right)^2 + \left(\frac{\delta u}{u}\right)^2 + \cdots + \left(\frac{\delta w}{w}\right)^2}$$

(provided all errors are independent and random)

and

$$\frac{\delta q}{|q|} \leq \frac{\delta x}{|x|} + \cdots + \frac{\delta z}{|z|} + \frac{\delta u}{|u|} + \cdots + \frac{\delta w}{|w|}$$

(always).

[See (3.18) & (3.19)]

Measured Quantity Times Exact Number: If B is known exactly and

$$q = Bx,$$

then

$$\delta q = |B| \delta x \quad \text{or, equivalently,} \quad \frac{\delta q}{|q|} = \frac{\delta x}{|x|}. \quad [\text{See (3.9)}]$$

Uncertainty in a Power: If n is an exact number and

$$q = x^n,$$

then

$$\frac{\delta q}{|q|} = |n| \frac{\delta x}{|x|}. \quad [\text{See (3.26)}]$$

Uncertainty in a Function of One Variable: If $q = q(x)$ is any function of x , then

$$\delta q = \left| \frac{dq}{dx} \right| \delta x. \quad [\text{See (3.23)}]$$

Sometimes, if $q(x)$ is complicated and if you have written a program to calculate $q(x)$ then, instead of differentiating $q(x)$, you may find it easier to use the equivalent

formula,

$$\delta q = |q(x_{\text{best}} + \delta x) - q(x_{\text{best}})|. \quad [\text{See Problem 3.32}]$$

General Formula for Error Propagation: If $q = q(x, \dots, z)$ is any function of x, \dots, z , then

$$\delta q = \sqrt{\left(\frac{\partial q}{\partial x} \delta x\right)^2 + \dots + \left(\frac{\partial q}{\partial z} \delta z\right)^2}$$

(provided all errors are independent and random)

and

$$\delta q \leq \left| \frac{\partial q}{\partial x} \right| \delta x + \dots + \left| \frac{\partial q}{\partial z} \right| \delta z$$

(always). [See (3.47) & (3.48)]

Problems for Chapter 3

For Section 3.2: The Square-Root Rule for a Counting Experiment

3.1.★ To measure the activity of a radioactive sample, two students count the alpha particles it emits. Student A watches for 3 minutes and counts 28 particles; Student B watches for 30 minutes and counts 310 particles. **(a)** What should Student A report for the average number emitted in 3 minutes, with his uncertainty? **(b)** What should Student B report for the average number emitted in 30 minutes, with her uncertainty? **(c)** What are the fractional uncertainties in the two measurements? Comment.

3.2.★ A nuclear physicist studies the particles ejected by a beam of radioactive nuclei. According to a proposed theory, the average rates at which particles are ejected in the forward and backward directions should be equal. To test this theory, he counts the total number ejected forward and backward in a certain 10-hour interval and finds 998 forward and 1,037 backward. **(a)** What are the uncertainties associated with these numbers? **(b)** Do these results cast any doubt on the theory that the average rates should be equal?

3.3.★ Most of the ideas of error analysis have important applications in many different fields. This applicability is especially true for the square-root rule (3.2) for counting experiments, as the following example illustrates. The normal average incidence of a certain kind of cancer has been established as 2 cases per 10,000 people per year. The suspicion has been aired that a certain town (population 20,000) suffers a high incidence of this cancer because of a nearby chemical dump. To test this claim, a reporter investigates the town's records for the past 4 years and finds 20 cases of the cancer. He calculates that the expected number is 16 (check this) and concludes that the observed rate is 25% more than expected. Is he justified in claiming that this result proves that the town has a higher than normal rate for this cancer?

3.4. ★★ As a sample of radioactive atoms decays, the number of atoms steadily diminishes and the sample's radioactivity decreases in proportion. To study this effect, a nuclear physicist monitors the particles ejected by a radioactive sample for 2 hours. She counts the number of particles emitted in a 1-minute period and repeats the measurement at half-hour intervals, with the following results:

Time elapsed, t (hours):	0.0	0.5	1.0	1.5	2.0
Number counted, ν , in 1 min:	214	134	101	61	54

- (a) Plot the number counted against elapsed time, including error bars to show the uncertainty in the numbers. (Neglect any uncertainty in the elapsed time.)
 (b) Theory predicts that the number of emitted particles should diminish exponentially as $\nu = \nu_0 \exp(-rt)$, where (in this case) $\nu_0 = 200$ and $r = 0.693 \text{ h}^{-1}$. On the same graph, plot this expected curve and comment on how well the data seem to fit the theoretical prediction.

For Section 3.3: Sums and Differences; Products and Quotients

3.5. ★ Using the provisional rules (3.4) and (3.8), compute the following:

- (a) $(5 \pm 1) + (8 \pm 2) - (10 \pm 4)$
- (b) $(5 \pm 1) \times (8 \pm 2)$
- (c) $(10 \pm 1)/(20 \pm 2)$
- (d) $(30 \pm 1) \times (50 \pm 1)/(5.0 \pm 0.1)$

3.6. ★ Using the provisional rules (3.4) and (3.8), compute the following:

- (a) $(3.5 \pm 0.1) + (8.0 \pm 0.2) - (5.0 \pm 0.4)$
- (b) $(3.5 \pm 0.1) \times (8.0 \pm 0.2)$
- (c) $(8.0 \pm 0.2)/(5.0 \pm 0.4)$
- (d) $(3.5 \pm 0.1) \times (8.0 \pm 0.2)/(5.0 \pm 0.4)$

3.7. ★ A student makes the following measurements:

$$\begin{aligned} a &= 5 \pm 1 \text{ cm}, & b &= 18 \pm 2 \text{ cm}, & c &= 12 \pm 1 \text{ cm}, \\ t &= 3.0 \pm 0.5 \text{ s}, & m &= 18 \pm 1 \text{ gram} \end{aligned}$$

Using the provisional rules (3.4) and (3.8), compute the following quantities with their uncertainties and percentage uncertainties: (a) $a + b + c$, (b) $a + b - c$, (c) ct , and (d) mb/t .

3.8. ★★ The binomial theorem states that for any number n and any x with $|x| < 1$,

$$(1 + x)^n = 1 + nx + \frac{n(n - 1)}{1 \cdot 2} x^2 + \frac{n(n - 1)(n - 2)}{1 \cdot 2 \cdot 3} x^3 + \dots$$

- (a) Show that if n is a positive integer, this infinite series terminates (that is, has only a finite number of nonzero terms). Write the series down explicitly for the cases $n = 2$ and $n = 3$. (b) Write down the binomial series for the case $n = -1$. This case gives an infinite series for $1/(1 + x)$, but when x is small, you get a good approximation if you keep just the first two terms:

$$\frac{1}{1 + x} \approx 1 - x,$$

as quoted in (3.6). Calculate both sides of this approximation for each of the values $x = 0.5, 0.1$, and 0.01 , and in each case find the percentage by which the approximation $(1 - x)$ differs from the exact value of $1/(1 + x)$.

For Section 3.4: Two Important Special Cases

3.9. ★ I measure the diameter of a circular disc as $d = 6.0 \pm 0.1$ cm and use this value to calculate the circumference $c = \pi d$ and radius $r = d/2$. What are my answers? [The rule (3.9) for “measured quantity \times exact number” applies to both of these calculations. In particular, you can write r as $d \times 1/2$, where the number $1/2$ is, of course, exact.]

3.10. ★ I have a set of callipers that can measure thicknesses of a few inches with an uncertainty of ± 0.005 inches. I measure the thickness of a deck of 52 cards and get 0.590 in. (a) If I now calculate the thickness of 1 card, what is my answer (including its uncertainty)? (b) I can improve this result by measuring several decks together. If I want to know the thickness of 1 card with an uncertainty of only 0.00002 in, how many decks do I need to measure together?

3.11. ★ With a good stopwatch and some practice, you can measure times ranging from approximately 1 second up to many minutes with an uncertainty of 0.1 second or so. Suppose that we wish to find the period τ of a pendulum with $\tau \approx 0.5$ s. If we time 1 oscillation, we have an uncertainty of approximately 20%; but by timing several oscillations together, we can do much better, as the following questions illustrate:

(a) If we measure the total time for 5 oscillations and get 2.4 ± 0.1 s, what is our final answer for τ , with its absolute and percent uncertainties? [Remember the rule (3.9).]

(b) What if we measure 20 oscillations and get 9.4 ± 0.1 s?

(c) Could the uncertainty in τ be improved indefinitely by timing more oscillations?

3.12. ★ If x has been measured as 4.0 ± 0.1 cm, what should I report for x^2 and x^3 ? Give percent and absolute uncertainties, as determined by the rule (3.10) for a power.

3.13. ★ If I have measured the radius of a sphere as $r = 2.0 \pm 0.1$ m, what should I report for the sphere’s volume?

3.14. ★ A visitor to a medieval castle measures the depth of a well by dropping a stone and timing its fall. She finds the time to fall is $t = 3.0 \pm 0.5$ sec and calculates the depth as $d = \frac{1}{2}gt^2$. What is her conclusion, if she takes $g = 9.80$ m/s² with negligible uncertainty?

3.15. ★★ Two students are asked to measure the rate of emission of alpha particles from a certain radioactive sample. Student A watches for 2 minutes and counts 32 particles. Student B watches for 1 hour and counts 786 particles. (The sample decays slowly enough that the expected rate of emission can be assumed to be constant during the measurements.) (a) What is the uncertainty in Student A’s result, 32, for the number of particles emitted in 2 minutes? (b) What is the uncertainty in Student B’s result, 786, for the number of particles emitted in 1 hour? (c) Each student now

divides his count by his number of minutes to find the *rate* of emission in particles per minute. Assuming the times, 2 min and 60 min, have negligible uncertainty, what are the two students' answers for the rate, with their uncertainties? Comment.

For Section 3.5: Independent Uncertainties in a Sum

3.16. ★ A student measures five lengths:

$$a = 50 \pm 5, \quad b = 30 \pm 3, \quad c = 60 \pm 2, \quad d = 40 \pm 1, \quad e = 5.8 \pm 0.3$$

(all in cm) and calculates the four sums $a + b$, $a + c$, $a + d$, $a + e$. Assuming the original errors were independent and random, find the uncertainties in her four answers [rule (3.13), “errors add in quadrature”]. If she has reason to think the original errors were *not* independent, what would she have to give for her final uncertainties [rule (3.14), “errors add directly”]? Assuming the uncertainties are needed with only one significant figure, identify those cases in which the second uncertainty (that in b , c , d , e) can be entirely ignored. If you decide to do the additions in quadrature on a calculator, note that the conversion from rectangular to polar coordinates automatically calculates $\sqrt{x^2 + y^2}$ for given x and y .

3.17. ★ Evaluate each of the following:

- (a) $(5.6 \pm 0.7) + (3.70 \pm 0.03)$
- (b) $(5.6 \pm 0.7) + (2.3 \pm 0.1)$
- (c) $(5.6 \pm 0.7) + (4.1 \pm 0.2)$
- (d) $(5.6 \pm 0.7) + (1.9 \pm 0.3)$

For each sum, consider both the case that the original uncertainties are independent and random (“errors add in quadrature”) and that they are not (“errors add directly”). Assuming the uncertainties are needed with only one significant figure, identify those cases in which the second of the original uncertainties can be ignored entirely. If you decide to do the additions in quadrature on a calculator, note that the conversion from rectangular to polar coordinates automatically calculates $\sqrt{x^2 + y^2}$ for given x and y .

For Section 3.6: More About Independent Uncertainties

3.18. ★ If you have not yet done it, do Problem 3.7 (assuming that the original uncertainties are *not* independent), and repeat each calculation assuming that the original uncertainties *are* independent and random. Arrange your answers in a table so that you can compare the two different methods of propagating errors.

3.19. ★ If you have not yet done it, do Problem 3.5 (assuming that the original uncertainties are *not* independent) and repeat each calculation assuming that the original uncertainties *are* independent and random. Arrange your answers in a table so that you can compare the two different methods of propagating errors.

3.20. ★ If you have not yet done it, do Problem 3.6 (assuming that the original uncertainties are *not* independent) and repeat each calculation assuming that the original uncertainties *are* independent and random. Arrange your answers in a table so that you can compare the two different methods of propagating errors.

3.21. ★ (a) To find the velocity of a cart on a horizontal air track, a student measures the distance d it travels and the time taken t as

$$d = 5.10 \pm 0.01 \text{ m} \quad \text{and} \quad t = 6.02 \pm 0.02 \text{ s.}$$

What is his result for $v = d/t$, with its uncertainty? (b) If he measures the cart's mass as $m = 0.711 \pm 0.002 \text{ kg}$, what would be his answer for the momentum $p = mv = md/t$? (Assume all errors are random and independent.)

3.22. ★ A student is studying the properties of a resistor. She measures the current flowing through the resistor and the voltage across it as

$$I = 2.10 \pm 0.02 \text{ amps} \quad \text{and} \quad V = 1.02 \pm 0.01 \text{ volts.}$$

(a) What should be her calculated value for the power delivered to the resistor, $P = IV$, with its uncertainty? (b) What for the resistance $R = V/I$? (Assume the original uncertainties are independent. With I in amps and V in volts, the power P comes out in watts and the resistance R in ohms.)

3.23. ★ In an experiment on the conservation of angular momentum, a student needs to find the angular momentum L of a uniform disc of mass M and radius R as it rotates with angular velocity ω . She makes the following measurements:

$$\begin{aligned} M &= 1.10 \pm 0.01 \text{ kg}, \\ R &= 0.250 \pm 0.005 \text{ m}, \\ \omega &= 21.5 \pm 0.4 \text{ rad/s} \end{aligned}$$

and then calculates L as $L = \frac{1}{2}MR^2\omega$. (The factor $\frac{1}{2}MR^2$ is just the moment of inertia of the uniform disc.) What is her answer for L with its uncertainty? (Consider the three original uncertainties independent and remember that the fractional uncertainty in R^2 is twice that in R .)

3.24. ★★ In his famous experiment with electrons, J.J. Thomson measured the "charge-to-mass ratio" $r = e/m$, where e is the electron's charge and m its mass. A modern classroom version of this experiment finds the ratio r by accelerating electrons through a voltage V and then bending them in a magnetic field. The ratio $r = e/m$ is given by the formula

$$r = \frac{125}{32\mu_0^2 N^2} \frac{D^2 V}{d^2 I^2}. \quad (3.53)$$

In this equation, μ_0 is the permeability constant of the vacuum (equal to $4\pi \times 10^{-7} \text{ N/A}^2$ exactly) and N is the number of turns in the coil that produces the magnetic field; D is the diameter of the field coils, V is the voltage that accelerates the electrons, d is the diameter of the electrons' curved path, and I is the current in the field coils. A student makes the following measurements:

$$\begin{aligned} N &= 72 \text{ (exactly)} \\ D &= 661 \pm 2 \text{ mm} \\ V &= 45.0 \pm 0.2 \text{ volts} \\ d &= 91.4 \pm 0.5 \text{ mm} \\ I &= 2.48 \pm 0.04 \text{ amps} \end{aligned}$$

(a) Find the student's answer for the charge-to-mass ratio of the electron, with its uncertainty. [Assume all uncertainties are independent and random. Note that the first factor in (3.53) is known exactly and can thus be treated as a single known constant, K . The second factor is a product and quotient of four numbers, D^2 , V , d^2 , and I^2 , so the fractional uncertainty in the final answer is given by the rule (3.18). Remember that the fractional uncertainty in D^2 is twice that in D , and so on.] (b) How well does this answer agree with the accepted value $r = 1.759 \times 10^{11} \text{ C/kg}$? (Note that you don't actually need to understand the theory of this experiment to do the problem. Nor do you need to worry about the units; if you use SI units for all the input quantities, the answer automatically comes out in the units given.)

3.25. ★★ We know from the rule (3.10) for uncertainties in a power that if $q = x^2$, the fractional uncertainty in q is twice that in x ;

$$\frac{\delta q}{q} = 2 \frac{\delta x}{|x|}.$$

Consider the following (fallacious) argument. We can regard x^2 as x times x ; so

$$q = x \times x;$$

therefore, by the rule (3.18),

$$\frac{\delta q}{q} = \sqrt{\left(\frac{\delta x}{x}\right)^2 + \left(\frac{\delta x}{x}\right)^2} = \sqrt{2} \frac{\delta x}{|x|}.$$

This conclusion is wrong. In a few sentences, explain why.

For Section 3.7: Arbitrary Functions of One Variable

3.26. ★ In nuclear physics, the energy of a subatomic article can be measured in various ways. One way is to measure how quickly the particle is stopped by an obstacle such as a piece of lead and then to use published graphs of energy versus stopping rate. Figure 3.7 shows such a graph for photons (the particles of light) in lead. The vertical axis shows the photons' energy E in MeV (millions of electron volts), and the horizontal axis shows the corresponding absorption coefficient μ in cm^2/g . (The precise definition of this coefficient need not concern us here; μ is simply a suitable measure of how quickly the photon is stopped in the lead.) From this graph, you can obviously find the energy E of a photon as soon as you know its absorption coefficient μ .

(a) A student observes a beam of photons (all with the same energy, E) and finds that their absorption coefficient in lead is $\mu = 0.10 \pm 0.01 \text{ cm}^2/\text{gram}$. Using the graph, find the energy E and the uncertainty δE . (You may find it helpful to draw on the graph the lines connecting the various points of interest, as done in Figure 3.3.) (b) What answer would the student have found if he had measured $\mu = 0.22 \pm 0.01 \text{ cm}^2/\text{gram}$?

3.27. ★ A student finds the refractive index n of a piece of glass by measuring the critical angle θ for light passing from the glass into air as $\theta = 41 \pm 1^\circ$. The relation

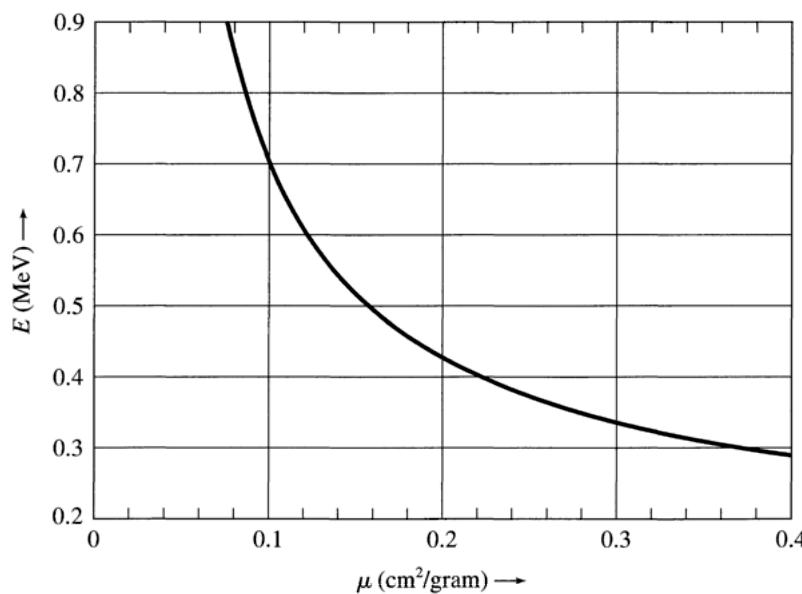


Figure 3.7. Energy E against absorption coefficient μ for photons in lead; for Problem 3.26.

between these is known to be $n = 1/\sin \theta$. Find the student's answer for n and use the rule (3.23) to find its uncertainty. (Don't forget to express $\delta\theta$ in radians.)

3.28. ★ (a) According to theory, the period T of a simple pendulum is $T = 2\pi\sqrt{L/g}$, where L is the length of the pendulum. If L is measured as $L = 1.40 \pm 0.01$ m, what is the predicted value of T ? **(b)** Would you say that a measured value of $T = 2.39 \pm 0.01$ s is consistent with the theoretical prediction of part (a)?

3.29. ★ (a) An experiment to measure Planck's constant h gives it in the form $h = K\lambda^{1/3}$ where K is a constant known exactly and λ is the measured wavelength emitted by a hydrogen lamp. If a student has measured λ with a fractional uncertainty she estimates as 0.3%, what will be the fractional uncertainty in her answer for h ? Comment. **(b)** If the student's best estimate for h is 6.644×10^{-34} J·s, is her result in satisfactory agreement with the accepted value of 6.626×10^{-34} J·s?

3.30. ★★ A spectrometer is a device for separating the different wavelengths in a beam of light and measuring the wavelengths. It deflects the different wavelengths through different angles θ , and, if the relation between the angle θ and wavelength λ is known, the experimenter can find λ by measuring θ . Careful measurements with a certain spectrometer have established the calibration curve shown in Figure 3.8; this figure is simply a graph of λ (in nanometers, or nm) against θ , obtained by measuring θ for several accurately known wavelengths λ . A student directs a narrow beam of light from a hydrogen lamp through this spectrometer and finds that the

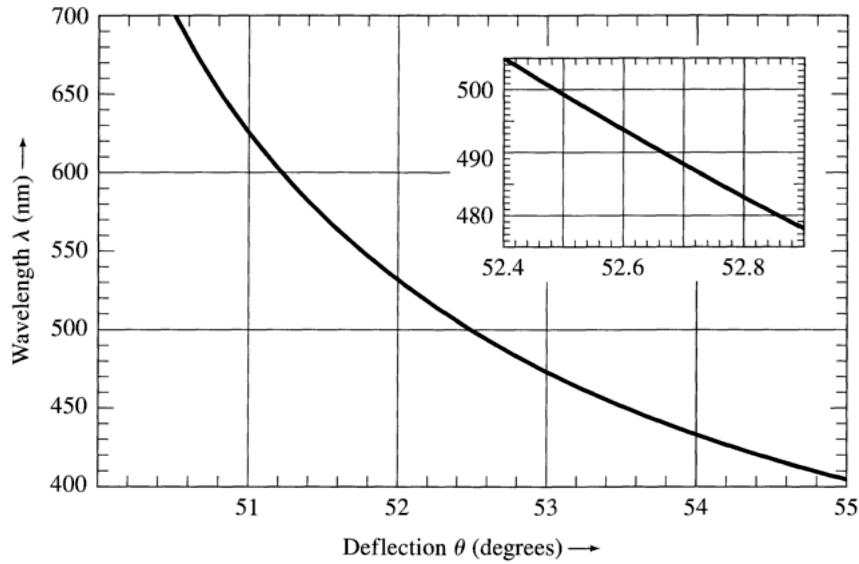


Figure 3.8. Calibration curve of wavelength λ against deflection θ for a spectrometer; for Problem 3.30.

light consists of just three well-defined wavelengths; that is, he sees three narrow beams (one red, one turquoise, and one violet) emerging at three different angles. He measures these angles as

$$\begin{aligned}\theta_1 &= 51.0 \pm 0.2^\circ \\ \theta_2 &= 52.6 \pm 0.2^\circ \\ \theta_3 &= 54.0 \pm 0.2^\circ\end{aligned}$$

(a) Use the calibration curve of Figure 3.8 to find the corresponding wavelengths λ_1 , λ_2 , and λ_3 with their uncertainties. (b) According to theory, these wavelengths should be 656, 486, and 434 nm. Are the student's measurements in satisfactory agreement with these theoretical values? (c) If the spectrometer has a vernier scale to read the angles, the angles can be measured with an uncertainty of 0.05° or even less. Let us suppose the three measurements above have uncertainties of $\pm 0.05^\circ$. Given this new, smaller uncertainty in the angles and *without drawing any more lines on the graph*, use your answers from part (a) to find the new uncertainties in the three wavelengths, explaining clearly how you do it. (Hint: the calibration curve is nearly straight in the vicinity of any one measurement.) (d) To take advantage of more accurate measurements, an experimenter may need to enlarge the calibration curve. The inset in Figure 3.8 is an enlargement of the vicinity of the angle θ_2 . Use this graph to find the wavelength λ_2 if θ_2 has been measured as $52.72 \pm 0.05^\circ$; check that your prediction for the uncertainty of λ_2 in part (c) was correct.

3.31. ★★ (a) An angle θ is measured as $125 \pm 2^\circ$, and this value is used to compute $\sin \theta$. Using the rule (3.23), calculate $\sin \theta$ and its uncertainty. (b) If a is measured as $a_{\text{best}} \pm \delta a$, and this value used to compute $f(a) = e^a$, what are f_{best} and

δf ? If $a = 3.0 \pm 0.1$, what are e^a and its uncertainty? (c) Repeat part (b) for the function $f(a) = \ln a$.

3.32. ★★★ The rule (3.23), $\delta q = |dq/dx|\delta x$, usually allows the uncertainty in a function $q(x)$ to be found quickly and easily. Occasionally, if $q(x)$ is very complicated, evaluating its derivative may be a nuisance, and going back to (3.20), from which (3.23) was derived, is sometimes easier. Note, however, that (3.20) was derived for a function whose slope was positive; if the slope is negative, the signs need to be reversed, and the general form of (3.20) is

$$\delta q = |q(x_{\text{best}} + \delta x) - q(x_{\text{best}})|. \quad (3.54)$$

Particularly if you have programmed your calculator or computer to find $q(x)$, then finding $q(x_{\text{best}} + \delta x)$ and $q(x_{\text{best}})$ and their difference will be easy.

(a) If you have a computer or programmable calculator, write a program to calculate the function

$$q(x) = \frac{(1 + x^2)^3}{x^2 + \cot x}.$$

Use this program to find $q(x)$ if $x = 0.75 \pm 0.1$, using the new rule (3.54) to find δq . (b) If you have the courage, differentiate $q(x)$ and check your value of δq using the rule (3.23).

3.33. ★★★ Do Problem 3.32 but use the function

$$q(x) = (1 - x^2) \cos\left(\frac{x + 2}{x^3}\right)$$

and the measured value $x = 1.70 \pm 0.02$.

For Section 3.8: Propagation Step by Step

3.34 ★ Use step-by-step propagation to find the following quantities (assuming that all given uncertainties are independent and random):

(a) $(20 \pm 1) + [(5.0 \pm 0.4) \times (3.0 \pm 0.2)]$

(b) $(20 \pm 1)/[(5.0 \pm 0.1) - (3.0 \pm 0.1)]$

(c) $(1.5 \pm 0.1) - 2 \sin(30 \pm 6^\circ)$

[In part (c), the number 2 is exact.]

3.35. ★ Use step-by-step propagation to find the following quantities (assuming that all given uncertainties are independent and random):

(a) $(20 \pm 1) + [(50 \pm 1)/(5.0 \pm 0.2)]$

(b) $(20 \pm 1) \times [(30 \pm 1) - (24 \pm 1)]$

(c) $(2.0 \pm 0.1) \times \tan(45 \pm 3^\circ)$

3.36. ★ Calculate the following quantities in steps as described in Section 3.8.

Assume all uncertainties are independent and random.

(a) $(12 \pm 1) \times [(25 \pm 3) - (10 \pm 1)]$

(b) $\sqrt{16 \pm 4} + (3.0 \pm 0.1)^3(2.0 \pm 0.1)$

(c) $(20 \pm 2)e^{-(1.0 \pm 0.1)}$

3.37. ★ (a) To find the acceleration of a glider moving down a sloping air track, I measure its velocities (v_1 and v_2) at two points and the time t it takes between them, as follows:

$$v_1 = 0.21 \pm 0.05, \quad v_2 = 0.85 \pm 0.05$$

(both in m/s) and

$$t = 8.0 \pm 0.1 \text{ s.}$$

Assuming all uncertainties are independent and random, what should I report for the acceleration, $a = (v_2 - v_1)/t$ and its uncertainty? **(b)** I have calculated theoretically that the acceleration should be $0.13 \pm 0.01 \text{ m/s}^2$. Does my measurement agree with this prediction?

3.38. ★ (a) As in Problem 3.37, I measure the velocities, v_1 and v_2 , of a glider at two points on a sloping air track with the results given there. Instead of measuring the time between the two points, I measure the distance as

$$d = 3.740 \pm 0.002 \text{ m.}$$

If I now calculate the acceleration as $a = (v_2^2 - v_1^2)/2d$, what should be my answer with its uncertainty? **(b)** How well does it agree with my theoretical prediction that $a = 0.13 \pm 0.01 \text{ m/s}^2$?

3.39. ★★ (a) The glider on a horizontal air track is attached to a spring that causes it to oscillate back and forth. The total energy of the system is $E = \frac{1}{2}mv^2 + \frac{1}{2}kx^2$, where m is the glider's mass, v is its velocity, k is the spring's force constant, and x is the extension of the spring from equilibrium. A student makes the following measurements:

$$\begin{aligned} m &= 0.230 \pm 0.001 \text{ kg}, & v &= 0.89 \pm 0.01 \text{ m/s}, \\ k &= 1.03 \pm 0.01 \text{ N/m}, & x &= 0.551 \pm 0.005 \text{ m}. \end{aligned}$$

What is her answer for the total energy E ? **(b)** She next measures the position x_{\max} of the glider at the extreme end of its oscillation, where $v = 0$, as

$$x_{\max} = 0.698 \pm 0.002 \text{ m.}$$

What is her value for the energy at the end point? **(c)** Are her results consistent with conservation of energy, which requires that these two energies should be the same?

For Section 3.9: Examples

3.40. ★★ Review the discussion of the simple pendulum in Section 3.9. In a real experiment, one should measure the period T for several different lengths l and hence obtain several different values of g for comparison. With a little thought, you can organize all data and calculations so that they appear in a single convenient tabulation, as in Table 3.2. Using Table 3.2 (or some other arrangement that you prefer), calculate g and its uncertainty δg for the four pairs of data shown. Are your answers consistent with the accepted value, 980 cm/s^2 ? Comment on the variation of δg as l gets smaller. (The answers given for the first pair of data will let you check your method of calculation.)

Table 3.2. Finding g with a pendulum; for Problem 3.40.

l (cm) all ± 0.1	T (sec) all ± 0.001	g (cm/s 2)	$\delta l/l$ (%)	$\delta T/T$ (%)	$\delta g/g$ (%)	answer $g \pm \delta g$
93.8	1.944	980	0.1	0.05	0.14	980 \pm 1.4
70.3	1.681					
45.7	1.358					
21.2	0.922					

3.41. ★★ Review the measurement of the refractive index of glass in Section 3.9. Using a table similar to Table 3.1, calculate the refractive index n and its fractional uncertainty for the data in Table 3.3. Are your answers consistent with the manufacturer's claim that $n = 1.50$? Comment on the variation in the uncertainties. (All angles are in degrees, i is the angle of incidence, r that of refraction.)

Table 3.3. Refractive index data (in degrees); for Problem 3.41.

i (all ± 1)	10	20	30	50	70
r (all ± 1)	7	13	20	29	38

For Section 3.10: A More Complicated Example

3.42. ★★★ Review the experiment in Section 3.10, in which a cart is rolled down an incline of slope θ . (a) If the cart's wheels are smooth and light, the expected acceleration is $g \sin \theta$. If θ is measured as 5.4 ± 0.1 degrees, what are the expected acceleration and its uncertainty? (b) If the experiment is repeated giving the cart various pushes at the top of the slope, the data and all calculations can be recorded as usual, in a single tabulation like Table 3.4. Using Equation (3.33) for the acceler-

Table 3.4. Acceleration data; for Problem 3.42.

t_1 (s) all ± 0.001	t_2 (s) all ± 0.001	$\frac{1}{t_1^2}$	$\frac{1}{t_2^2}$	$\frac{1}{t_2^2} - \frac{1}{t_1^2}$	a (cm/s 2)
$0.054 \pm 2\%$	$0.031 \pm 3\%$	343 ± 14	1040 ± 62	698 ± 64	87 ± 8
0.038	0.027				
0.025	0.020				

ation (and the same values $l^2/2s = 0.125$ cm $\pm 2\%$ as before), calculate a and δa for the data shown. Are the results consistent with the expected constancy of a and with the expected value $g \sin \theta$ of part (a)? Would pushing the cart harder to check the constancy of a at even higher speeds be worthwhile? Explain.

For Section 3.11: General Formula for Error Propagation

3.43. ★ The partial derivative $\partial q/\partial x$ of $q(x, y)$ is obtained by differentiating q with respect to x while treating y as a constant. Write down the partial derivatives $\partial q/\partial x$ and $\partial q/\partial y$ for the three functions:

$$(a) q(x, y) = x + y, \quad (b) q(x, y) = xy, \quad (c) q(x, y) = x^2y^3.$$

3.44. ★★ The crucial approximation used in Section 3.11 relates the value of the function q at the point $(x + u, y + v)$ to that at the nearby point (x, y) :

$$q(x + u, y + v) \approx q(x, y) + \frac{\partial q}{\partial x}u + \frac{\partial q}{\partial y}v \quad (3.55)$$

when u and v are small. Verify explicitly that this approximation is good for the three functions of Problem 3.43. That is, for each function, write both sides of Equation (3.55) exactly, and show that they are approximately equal when u and v are small. For example, if $q(x, y) = xy$, then the left side of Equation (3.55) is

$$(x + u)(y + v) = xy + uy + xv + uv.$$

As you will show, the right side of (3.55) is

$$xy + yu + xv.$$

If u and v are small, then uv can be neglected in the first expression, and the two expressions are approximately equal.

3.45. ★ (a) For the function $q(x, y) = xy$, write the partial derivatives $\partial q/\partial x$ and $\partial q/\partial y$. Suppose we measure x and y with uncertainties δx and δy and then calculate $q(x, y)$. Using the general rules (3.47) and (3.48), write the uncertainty δq both for the case when δx and δy are independent and random, and for the case when they are not. Divide through by $|q| = |xy|$, and show that you recover the simple rules (3.18) and (3.19) for the fractional uncertainty in a product. **(b)** Repeat part (a) for the function $q(x, y) = x^n y^m$, where n and m are known fixed numbers. **(c)** What do Equations (3.47) and (3.48) become when $q(x)$ depends on only one variable?

3.46. ★★ If you measure two independent variables as

$$x = 6.0 \pm 0.1 \quad \text{and} \quad y = 3.0 \pm 0.1,$$

and use these values to calculate $q = xy + x^2/y$, what will be your answer and its uncertainty? [You must use the general rule (3.47) to find δq . To simplify your calculation, do it by first finding the two separate contributions δq_x and δq_y as defined in (3.50) and (3.51) and then combining them in quadrature.]

3.47. ★★ The Atwood machine consists of two masses M and m (with $M > m$) attached to the ends of a light string that passes over a light, frictionless pulley. When the masses are released, the mass M is easily shown to accelerate down with an acceleration

$$a = g \frac{M - m}{M + m}.$$

Suppose that M and m are measured as $M = 100 \pm 1$ and $m = 50 \pm 1$, both in grams. Use the general rule (3.47) to derive a formula for the uncertainty in the expected acceleration δa in terms of the masses and their uncertainties and then find δa for the given numbers.

3.48. ★★★ If we measure three independent quantities x , y , and z and then calculate a function such as $q = (x + y)/(x + z)$, then, as discussed at the beginning of Section 3.11, a stepwise calculation of the uncertainty in q may overestimate the uncertainty δq . (a) Consider the measured values $x = 20 \pm 1$, $y = 2$, and $z = 0$, and for simplicity, suppose that δy and δz are negligible. Calculate the uncertainty δq correctly using the general rule (3.47) and compare your result with what you would get if you were to calculate δq in steps. (b) Do the same for the values $x = 20 \pm 1$, $y = -40$, and $z = 0$. Explain any differences between parts (a) and (b).

3.49. ★★★ If an object is placed at a distance p from a lens and an image is formed at a distance q from the lens, the lens's focal length can be found as

$$f = \frac{pq}{p + q}. \quad (3.56)$$

[This equation follows from the “lens equation,” $1/f = (1/p) + (1/q)$.] (a) Use the general rule (3.47) to derive a formula for the uncertainty δf in terms of p , q , and their uncertainties. (b) Starting from (3.56) directly, you cannot find δf in steps because p and q both appear in numerator and denominator. Show, however, that f can be rewritten as

$$f = \frac{1}{(1/p) + (1/q)}.$$

Starting from this form, you *can* evaluate δf in steps. Do so, and verify that you get the same answer as in part (a).

3.50. ★★★ Suppose you measure three independent variables as

$$x = 10 \pm 2, \quad y = 7 \pm 1, \quad \theta = 40 \pm 3^\circ,$$

and use these values to compute

$$q = \frac{x + 2}{x + y \cos(4\theta)}.$$

What should be your answer for q and its uncertainty? Note that you cannot do the error propagation in steps here because the variable x appears in both numerator and denominator; therefore, you must use the general rule (3.47).

Chapter 4

Statistical Analysis of Random Uncertainties

We have seen that one of the best ways to assess the reliability of a measurement is to repeat it several times and examine the different values obtained. In this chapter and Chapter 5, I describe statistical methods for analyzing measurements in this way.

As noted before, not all types of experimental uncertainty can be assessed by statistical analysis based on repeated measurements. For this reason, uncertainties are classified into two groups: the *random* uncertainties, which *can* be treated statistically, and the *systematic* uncertainties, which *cannot*. This distinction is described in Section 4.1. Most of the remainder of this chapter is devoted to random uncertainties. Section 4.2 introduces, without formal justification, two important definitions related to a series of measured values x_1, \dots, x_N , all of some single quantity x . First, I define the *average* or *mean* \bar{x} of x_1, \dots, x_N . Under suitable conditions, \bar{x} is the best estimate of x based on the measured values x_1, \dots, x_N . I then define the *standard deviation* of x_1, \dots, x_N , which is denoted σ_x and characterizes the average uncertainty in the separate measured values x_1, \dots, x_N . Section 4.3 gives an example of the use of the standard deviation.

Section 4.4 introduces the important notion of the *standard deviation of the mean*. This parameter is denoted $\sigma_{\bar{x}}$ and characterizes the uncertainty in the mean \bar{x} as the best estimate for x . Section 4.5 gives examples of the standard deviation of the mean. Finally, in Section 4.6, I return to the vexing problem of systematic errors.

Nowhere in this chapter do I attempt a complete justification of the methods described. The main aim is to introduce the basic formulas and describe how they are used. In Chapter 5, I give proper justifications, based on the important idea of the normal distribution curve.

The relation of the material of this chapter (statistical analysis) to the material of Chapter 3 (error propagation) deserves mention. From a practical point of view, these two topics can be viewed as separate, though related, branches of error analysis (somewhat as algebra and geometry are separate, though related, branches of mathematics). Both topics need to be mastered, because most experiments require the use of both.

In a few kinds of experiments, the roles of error propagation and of statistical analysis are complementary. That is, the experiment can be analyzed using either

error propagation or statistical methods. Consider an example: Suppose you decide to measure the acceleration of gravity, g , by measuring the period, T , and the length, l , of a simple pendulum. Since $T = 2\pi\sqrt{l/g}$, you can find g as $g = 4\pi^2 l/T^2$. You might decide to repeat this experiment using several different values of l and measuring the corresponding period T for each. In this way, you would arrive at several values for g . To find the uncertainty in these values of g , you could proceed in either of two ways. If you can estimate realistically the uncertainties in your measurements of l and T , you could propagate these uncertainties to find the uncertainties in your values of g . Alternatively, given your several values of g , you could analyze them statistically; in particular, their *standard deviation* will be a good measure of their uncertainty. Unfortunately, you do not truly have a choice of how to find the uncertainty. If the uncertainty can be found in these two ways, you really ought to do so *both* ways to check that they do give, at least approximately, the same answer.

4.1 Random and Systematic Errors

Experimental uncertainties that can be revealed by repeating the measurements are called *random* errors; those that cannot be revealed in this way are called *systematic*. To illustrate this distinction, let us consider some examples. Suppose first that we time a revolution of a steadily rotating turntable. One source of error will be our reaction time in starting and stopping the watch. If our reaction time were always exactly the same, these two delays would cancel one another. In practice, however, our reaction time will vary. We may delay more in starting, and so underestimate the time of a revolution; or we may delay more in stopping, and so overestimate the time. Since either possibility is equally likely, the sign of the effect is *random*. If we repeat the measurement several times, we will sometimes overestimate and sometimes underestimate. Thus, our variable reaction time will show up as a variation of the answers found. By analyzing the spread in results statistically, we can get a very reliable estimate of this kind of error.

On the other hand, if our stopwatch is running consistently slow, then all our times will be underestimates, and no amount of repetition (with the same watch) will reveal this source of error. This kind of error is called *systematic*, because it always pushes our result in the same direction. (If the watch runs slow, we always underestimate; if the watch runs fast, we always overestimate.) Systematic errors cannot be discovered by the kind of statistical analysis contemplated here.

As a second example of random versus systematic errors, suppose we have to measure some well-defined length with a ruler. One source of uncertainty will be the need to interpolate between scale markings; and this uncertainty is probably random. (When interpolating, we are probably just as likely to overestimate as to underestimate.) But there is also the possibility that our ruler has become distorted; and this source of uncertainty would probably be systematic. (If the ruler has stretched, we always underestimate; if it has shrunk, we always overestimate.)

Just as in these two examples, almost all measurements are subject to both random and systematic uncertainties. You should have no difficulty finding more examples. In particular, notice that common sources of random uncertainties are

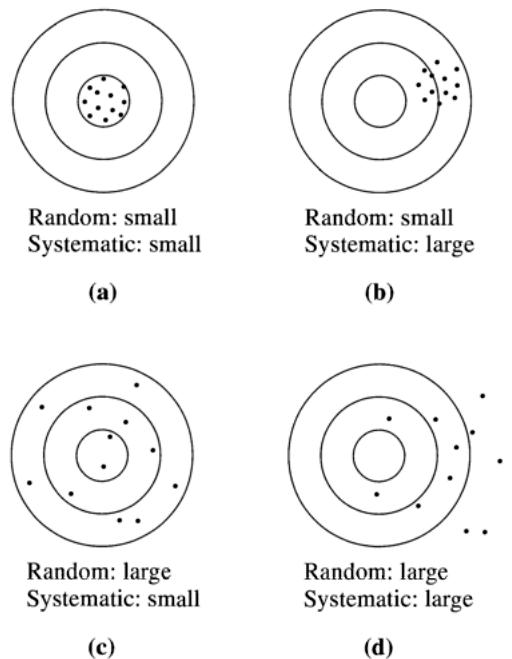


Figure 4.1. Random and systematic errors in target practice. **(a)** Because all shots arrived close to one another, we can tell the random errors are small. Because the distribution of shots is centered on the center of the target, the systematic errors are also small. **(b)** The random errors are still small, but the systematic ones are much larger—the shots are “systematically” off-center toward the right. **(c)** Here, the random errors are large, but the systematic ones are small—the shots are widely scattered but not systematically off-center. **(d)** Here, both random and systematic errors are large.

small errors of judgment by the observer (as when interpolating), small disturbances of the apparatus (such as mechanical vibrations), problems of definition, and several others. Perhaps the most obvious cause of systematic error is the miscalibration of instruments, such as the watch that runs slow, the ruler that has been stretched, or a meter that is improperly zeroed.

To get a better feel for the difference between random and systematic errors, consider the analogy shown in Figure 4.1. Here the “experiment” is a series of shots fired at a target; accurate “measurements” are shots that arrive close to the center. Random errors are caused by anything that makes the shots arrive at randomly different points. For example, the marksman may have an unsteady hand, or fluctuating atmospheric conditions between the marksman and the target may distort the view of the target in a random way. Systematic errors arise if anything makes the shots arrive off-center in one “systematic” direction, for instance, if the gun’s sights are misaligned. Note from Figure 4.1 how the results change according to the various combinations of small or large random or systematic errors.

Although Figure 4.1 is an excellent illustration of the effects of random and systematic errors, it is, nonetheless, misleading in one important respect. Because

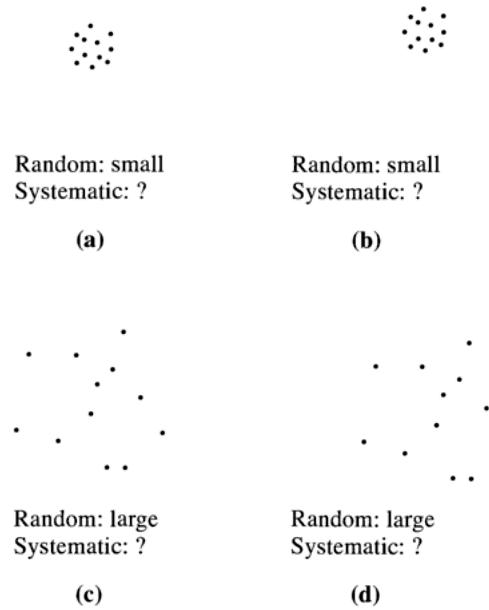


Figure 4.2. The same experiment as in Figure 4.1 redrawn without showing the position of the target. This situation corresponds closely to the one in most real experiments, in which we do not know the true value of the quantity being measured. Here, we can still assess the random errors easily but cannot tell anything about the systematic ones.

each of the four pictures shows the position of the target, we can tell at a glance whether a particular shot was accurate or not. In particular, the difference between the top two pictures is immediately evident. The shots in the left picture cluster around the target's center, whereas those in the right picture cluster around a point well off-center; clearly, therefore, the marksman responsible for the left picture had little systematic error, but the one responsible for the right picture had a lot more. Knowing the position of the target in Figure 4.1 corresponds, in a laboratory measurement, to knowing the true value of the measured quantity, and in the vast majority of real measurements, we do *not* know this true value. (If we knew the true value, we would usually not bother to measure it.)

To improve the analogy of Figure 4.1 with most real experiments, we need to redraw it without the rings that show the position of the target, as in Figure 4.2. In these pictures, identifying the random errors is still easy. (The top two pictures still obviously have smaller random errors than the bottom two.) Determining which marksman had larger systematic errors, however, is *impossible* based on Figure 4.2. This situation is exactly what prevails in most real experiments; by examining the distribution of measured values, we can easily assess the random errors but get no guidance concerning the systematic errors.

The distinction between random and systematic errors is not always clear-cut, and a problem that causes random errors in one experiment may produce systematic errors in another. For example, if you position your head first to one side and then to another to read a typical meter (such as an ordinary clock), the reading on the meter changes. This effect, called *parallax*, means that a meter can be read correctly only if you position yourself directly in front of it. No matter how careful you are, you cannot always position your eye *exactly* in front of the meter; consequently, your measurements will have a small uncertainty due to parallax, and this uncertainty will probably be random. On the other hand, a careless experimenter who places a meter to one side of his seat and forgets to worry about parallax will introduce a systematic error into all his readings. Thus, the same effect, parallax, can produce random uncertainties in one case, and systematic uncertainties in another.

The treatment of random errors is different from that of systematic errors. The statistical methods described in the following sections give a reliable estimate of the random uncertainties, and, as we shall see, provide a well-defined procedure for reducing them. For the reasons just discussed, systematic uncertainties are usually hard to evaluate and even to detect. The experienced scientist has to learn to anticipate the possible sources of systematic error and to make sure that all systematic errors are much less than the required precision. Doing so will involve, for example, checking the meters against accepted standards and correcting them or buying better ones if necessary. Unfortunately, in the first-year physics laboratory, such checks are rarely possible, so the treatment of systematic errors is often awkward. This concept is discussed further in Section 4.6. For now, I will discuss experiments in which all sources of systematic error have been identified and made much smaller than the required precision.

4.2 The Mean and Standard Deviation

Suppose we need to measure some quantity x , and we have identified all sources of systematic error and reduced them to a negligible level. Because all remaining sources of uncertainty are random, we should be able to detect them by repeating the measurement several times. We might, for example, make the measurement five times and find the results

$$71, 72, 72, 73, 71 \quad (4.1)$$

(where, for convenience, we have omitted any units).

The first question we address is this: Given the five measured values (4.1), what should we take for our best estimate x_{best} of the quantity x ? Reasonably, our best estimate would seem to be the *average* or *mean* \bar{x} of the five values found, and in Chapter 5, I will prove that this choice is normally best. Thus,

$$\begin{aligned} x_{\text{best}} &= \bar{x} \\ &= \frac{71 + 72 + 72 + 73 + 71}{5} \\ &= 71.8. \end{aligned} \quad (4.2)$$

Here, the second line is simply the definition of the mean \bar{x} for the numbers at hand.¹

More generally, suppose we make N measurements of the quantity x (all using the same equipment and procedures) and find the N values

$$x_1, x_2, \dots, x_N. \quad (4.3)$$

Once again, the best estimate for x is usually the average of x_1, \dots, x_N . That is,

$$x_{\text{best}} = \bar{x}, \quad (4.4)$$

where

$$\begin{aligned} \bar{x} &= \frac{x_1 + x_2 + \dots + x_N}{N} \\ &= \frac{\sum x_i}{N}. \end{aligned} \quad (4.5)$$

In the last line, I have introduced the useful sigma notation, according to which

$$\sum_{i=1}^N x_i = \sum_i x_i = \sum x_i = x_1 + x_2 + \dots + x_N;$$

the second and third expressions here are common abbreviations, which I will use when there is no danger of confusion.

The concept of the average or mean is almost certainly familiar to most readers. Our next concept, that of the *standard deviation*, is probably less so. The standard deviation of the measurements x_1, \dots, x_N is an estimate of the *average uncertainty of the measurements* x_1, \dots, x_N and is determined as follows.

Given that the mean \bar{x} is our best estimate of the quantity x , it is natural to consider the difference $x_i - \bar{x} = d_i$. This difference, often called the *deviation* (or residual) of x_i from \bar{x} , tells us *how much the i^{th} measurement x_i differs from the average \bar{x}* . If the deviations $d_i = x_i - \bar{x}$ are all very small, our measurements are all close together and presumably very precise. If some of the deviations are large, our measurements are obviously not so precise.

To be sure you understand the idea of the deviation, let us calculate the deviations for the set of five measurements reported in (4.1). These deviations can be listed as shown in Table 4.1. Notice that the deviations are not (of course) all the same size; d_i is small if the i^{th} measurement x_i happens to be close to \bar{x} , but d_i is large if x_i is far from \bar{x} . Notice also that some of the d_i are positive and some negative because some of the x_i are bound to be higher than the average \bar{x} , and some are bound to be lower.

To estimate the average reliability of the measurements x_1, \dots, x_5 , we might naturally try averaging the deviations d_i . Unfortunately, as a glance at Table 4.1 shows, the average of the deviations is zero. In fact, this average will be zero for

¹ In this age of pocket calculators, it is worth pointing out that an average such as (4.2) is easily calculated in your head. Because all the numbers are in the seventies, the same must be true of the average. All that remains is to average the numbers 1, 2, 2, 3, 1 in the units place. These numbers obviously average to $9/5 = 1.8$, and our answer is $\bar{x} = 71.8$.

Table 4.1. Calculation of deviations.

Trial number <i>i</i>	Measured value <i>x_i</i>	Deviation <i>d_i</i> = <i>x_i</i> - \bar{x}
1	71	-0.8
2	72	0.2
3	72	0.2
4	73	1.2
5	71	-0.8
$\Sigma x_i = 359$		$\Sigma d_i = 0.0$
mean, $\bar{x} = \Sigma x_i/N = 359/5 = 71.8$		

any set of measurements x_1, \dots, x_N because the definition of the average \bar{x} ensures that $d_i = x_i - \bar{x}$ is sometimes positive and sometimes negative in just such a way that \bar{d} is zero (see Problem 4.4). Obviously, then, the average of the deviations is not a useful way to characterize the reliability of the measurements x_1, \dots, x_N .

The best way to avoid this annoyance is to *square* all the deviations, which will create a set of *positive* numbers, and then average these numbers.² If we then take the square root of the result, we obtain a quantity with the same units as x itself. This number is called the *standard deviation* of x_1, \dots, x_N , and is denoted σ_x :

$$\sigma_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (d_i)^2} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}. \quad (4.6)$$

With this definition, the standard deviation can be described as the *root mean square* (or RMS) deviation of the measurements x_1, \dots, x_N . It proves to be a useful way to characterize the reliability of the measurements. [As we will discuss shortly, the definition (4.6) is sometimes modified by replacing the denominator N by $N - 1$.]

To calculate the standard deviation σ_x as defined by (4.6), we must compute the deviations d_i , square them, average these squares, and then take the square root of the result. For the data of Table 4.1, we start this calculation in Table 4.2.

Table 4.2. Calculation of the standard deviation.

Trial number <i>i</i>	Measured value <i>x_i</i>	Deviation <i>d_i</i> = <i>x_i</i> - \bar{x}	Deviation squared <i>d_i</i> ²
1	71	-0.8	0.64
2	72	0.2	0.04
3	72	0.2	0.04
4	73	1.2	1.44
5	71	-0.8	0.64
$\Sigma x_i = 359$		$\Sigma d_i = 0.0$	$\Sigma d_i^2 = 2.80$
$\bar{x} = 359/5 = 71.8$			

²Another possibility would be to take the absolute values $|d_i|$ and average them, but the average of the d_i^2 proves more useful. The average of the $|d_i|$ is sometimes (misleadingly) called the *average deviation*.

Summing the numbers d_i^2 in the fourth column of Table 4.2 and dividing by 5, we obtain the quantity σ_x^2 (often called the *variance* of the measurements),

$$\sigma_x^2 = \frac{1}{N} \sum d_i^2 = \frac{2.80}{5} = 0.56. \quad (4.7)$$

Taking the square root, we find the standard deviation

$$\sigma_x \approx 0.7. \quad (4.8)$$

Thus the average uncertainty of the five measurements 71, 72, 72, 73, 71 is approximately 0.7.

Unfortunately, the standard deviation has an alternative definition. There are theoretical arguments for replacing the factor N in (4.6) by $(N - 1)$ and defining the standard deviation σ_x of x_1, \dots, x_N as

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum d_i^2} = \sqrt{\frac{1}{N-1} \sum (x_i - \bar{x})^2}. \quad (4.9)$$

I will not try here to prove that definition (4.9) of σ_x is better than (4.6), except to say that the new “improved” definition is obviously a little larger than the old one (4.6) and that (4.9) corrects a tendency for (4.6) to underestimate the uncertainty in the measurements x_1, \dots, x_N , especially if the number of measurements N is small. This tendency can be understood by considering the extreme (and absurd) case that $N = 1$ (that is, we make only one measurement). Here, the average \bar{x} is equal to our one reading x_1 , and the one deviation is automatically zero. Therefore, the definition (4.6) gives the absurd result $\sigma_x = 0$. On the other hand, the definition (4.9) gives 0/0; that is, with definition (4.9), σ_x is undefined, which correctly reflects our total ignorance of the uncertainty after just one measurement. The definition (4.6) is sometimes called the *population standard deviation* and (4.9) the *sample standard deviation*.

The difference between the two definitions (4.6) and (4.9) is almost always numerically insignificant. You should always repeat a measurement many times (at least five, and preferably many more). Even if you make only five measurements ($N = 5$), the difference between $\sqrt{N} = 2.2$ and $\sqrt{N-1} = 2$ is, for most purposes, insignificant. For example, if we recalculate the standard deviation (4.8) using the improved definition (4.9), we obtain $\sigma_x = 0.8$ instead of $\sigma_x = 0.7$, not a very important difference. Nevertheless, you need to be aware of both definitions. In the physics laboratory, using the more conservative (that is, larger) definition (4.9) is almost always best, but in any case, your laboratory report should state clearly which definition you are using so that your readers can check the calculations for themselves.

Quick Check 4.1. You measure the time for a cart to roll down the same length of track four times and get the following results:

21, 24, 25, 22

(in seconds). Find the average time and the standard deviation as given by the improved definition (4.9).

To understand the notion of the standard deviation, you must be able to calculate it yourself for simple cases such as that in Quick Check 4.1. Most scientific calculators, however, have a built-in function to do the calculation automatically, and you will certainly want to use this function for real experiments that involve numerous measurements. If you are not sure how to use your calculator to obtain standard deviations, take the time to learn, and then use the function to check your answer to Quick Check 4.1. Some calculators give you a choice of the definitions (4.6) or (4.9); some use just (4.9). Make sure you know what yours does.

4.3 The Standard Deviation as the Uncertainty in a Single Measurement

Recall the claim that the standard deviation σ_x characterizes the average uncertainty of the measurements x_1, \dots, x_N from which it was calculated. In Chapter 5, I will justify this claim by proving the following more precise statement. If you measure the same quantity x many times, always using the same method, and if all your sources of uncertainty are small and random, then your results will be distributed around the true value x_{true} in accordance with the so-called normal, or bell-shaped, curve. In particular, *approximately 68% of your results³ will fall within a distance σ_x on either side of x_{true}* ; that is, 68% of your measurements will fall in the range $x_{\text{true}} \pm \sigma_x$.

In other words, if you make a *single* measurement (using the same method), the probability is 68% that your result will be within σ_x of the correct value. Thus, we can adopt σ_x to mean exactly what we have been calling “uncertainty.” If you make one measurement of x , the uncertainty associated with this measurement can be taken to be

$$\delta x = \sigma_x;$$

with this choice, you can be 68% confident that the measurement is within δx of the correct answer.

To illustrate the application of these ideas, suppose we are given a box of similar springs and told to measure their spring constants k . We might measure the spring constants by loading each spring and observing the resulting extension or, perhaps better, by suspending a mass from each spring and timing its oscillations. Whatever method we choose, we need to know k and its uncertainty δk for each spring, but it would be hopelessly time-consuming to repeat our measurements many times for each spring. Instead we reason as follows: If we measure k for the first

³As we will see, the exact number is 68.27 . . . %, but stating this kind of number so precisely is obviously absurd. In fact, it is often best to think of this number as “about two thirds.”

spring several (say, 5 or 10) times, then the mean of these measurements should give a good estimate of k for the first spring. More important for now, the standard deviation σ_k of these 5 or 10 measurements provides us with an estimate of the uncertainty in our method for measuring k . Provided our springs are all reasonably similar and we use the same method to measure each one, we can reasonably expect the same uncertainty in each measurement.⁴ Thus, for each subsequent spring we need to make only one measurement, and we can immediately state that the uncertainty δk is the standard deviation σ_k measured for the first spring, with a 68% confidence that our answer is within σ_k of the correct value.

To illustrate these ideas numerically, we can imagine making 10 measurements on the first spring and obtaining the following measured values of k (in newtons/meter):

$$86, 85, 84, 89, 85, 89, 87, 85, 82, 85. \quad (4.10)$$

From these values, we can immediately calculate $\bar{k} = 85.7$ N/m and, using the definition (4.9),

$$\sigma_k = 2.16 \text{ N/m} \quad (4.11)$$

$$\approx 2 \text{ N/m}. \quad (4.12)$$

The uncertainty in any one measurement of k is therefore approximately 2 N/m. If we now measure the second spring once and obtain the answer $k = 71$ N/m, we can without further ado take $\delta k = \sigma_k = 2$ N/m and state with 68% confidence that k lies in the range

$$(k \text{ for second spring}) = 71 \pm 2 \text{ N/m}. \quad (4.13)$$

4.4 The Standard Deviation of the Mean

If x_1, \dots, x_N are the results of N measurements of the same quantity x , then, as we have seen, our best estimate for the quantity x is their mean \bar{x} . We have also seen that the standard deviation σ_x characterizes the average uncertainty of the separate measurements x_1, \dots, x_N . Our answer $x_{\text{best}} = \bar{x}$, however, represents a judicious combination of all N measurements, and we have every reason to think it will be more reliable than any one of the measurements taken alone. In Chapter 5, I will prove that the uncertainty in the final answer $x_{\text{best}} = \bar{x}$ is given by the standard deviation σ_x divided by \sqrt{N} . This quantity is called the *standard deviation of the mean*, or SDOM, and is denoted $\sigma_{\bar{x}}$:

$$\sigma_{\bar{x}} = \sigma_x / \sqrt{N}. \quad (4.14)$$

(Other common names are *standard error* and *standard error of the mean*.) Thus, based on the N measured values x_1, \dots, x_N , we can state our final answer for the

⁴If some springs are very different from the first, our uncertainty in measuring them may be different. Thus, if the springs differ a lot, we would need to check our uncertainty by making several measurements for each of two or three different springs.

value of x as

$$(\text{value of } x) = x_{\text{best}} \pm \delta x,$$

where $x_{\text{best}} = \bar{x}$, the mean of x_1, \dots, x_N , and δx is the standard deviation of the mean,

$$\delta x = \sigma_{\bar{x}} = \sigma_x / \sqrt{N}. \quad (4.15)$$

As an example, we can consider the 10 measurements reported in (4.10) of the spring constant k of one spring. As we saw, the mean of these values is $\bar{k} = 85.7$ N/m, and the standard deviation is $\sigma_k = 2.2$ N/m. Therefore, the standard deviation of the mean is

$$\sigma_{\bar{k}} = \sigma_k / \sqrt{10} = 0.7 \text{ N/m}, \quad (4.16)$$

and our final answer, based on these 10 measurements, would be that the spring has

$$k = 85.7 \pm 0.7 \text{ newtons/meter.} \quad (4.17)$$

When you give an answer like this, you must state clearly what the numbers are—namely, the mean and the standard deviation of the mean—so your readers can judge their significance for themselves.

An important feature of the standard deviation of the mean, $\sigma_{\bar{x}} = \sigma_x / \sqrt{N}$, is the factor \sqrt{N} in the denominator. The standard deviation σ_x represents the average uncertainty in the individual measurements x_1, \dots, x_N . Thus, if we were to make some more measurements (using the same technique), the standard deviation σ_x would not change appreciably. On the other hand, the standard deviation of the mean, σ_x / \sqrt{N} , would slowly decrease as we increase N . This decrease is just what we would expect. If we make more measurements before computing an average, we would naturally expect the final result to be more reliable, and this improved reliability is just what the denominator \sqrt{N} in (4.15) guarantees. This conclusion provides one obvious way to improve the precision of our measurements.

Unfortunately, the factor \sqrt{N} grows rather slowly as we increase N . For example, if we wish to improve our precision by a factor of 10 simply by increasing the number of measurements N , we will have to increase N by a factor of 100—a daunting prospect, to say the least! Furthermore, we are for the moment neglecting systematic errors, and these are *not* reduced by increasing the number of measurements. Thus, in practice, if you want to increase your precision appreciably, you will probably do better to improve your technique than to rely merely on increased numbers of measurements.

Quick Check 4.2. A student makes five measurements of e , the magnitude of the electron's charge, as follows:

$$15; 17, 18, 14, 16,$$

all in units of 10^{-20} coulombs. Find her best estimate for e (as given by the mean) and its uncertainty (as given by the SDOM).

4.5 Examples

In this section, I discuss two examples of simple experiments that make use of the ideas of the past three sections.

Example: Area of a Rectangle

As a first, simple application of the standard deviation of the mean, imagine that we have to measure very accurately the area A of a rectangular plate approximately 2.5 cm \times 5 cm. We first find the best available measuring device, which might be a vernier caliper, and then make several measurements of the length l and breadth b of the plate. To allow for irregularities in the sides, we make our measurements at several different positions, and to allow for small defects in the instrument, we use several different calipers (if available). We might make 10 measurements each of l and b and obtain the results shown in Table 4.3.

Table 4.3. Length and breadth (in mm).

	Measured values	Mean	SD	SDOM
l	24.25, 24.26, 24.22, 24.28, 24.24 24.25, 24.22, 24.26, 24.23, 24.24	$\bar{l} = 24.245$	$\sigma_l = 0.019$	$\sigma_{\bar{l}} = 0.006$
b	50.36, 50.35, 50.41, 50.37, 50.36 50.32, 50.39, 50.38, 50.36, 50.38	$\bar{b} = 50.368$	$\sigma_b = 0.024$	$\sigma_{\bar{b}} = 0.008$

Using the 10 observed values of l , you can quickly calculate the mean \bar{l} , the standard deviation σ_l , and the standard deviation of the mean $\sigma_{\bar{l}}$, as shown in the columns labeled mean, SD, and SDOM. In the same way you can calculate \bar{b} , σ_b , and $\sigma_{\bar{b}}$. Before doing any further calculations, you should examine these results to see if they seem reasonable. For example, the two standard deviations σ_l and σ_b are supposed to be the average uncertainty in the measurements of l and b . Because l and b were measured in exactly the same way, σ_l and σ_b should not differ significantly from each other or from what we judge to be a reasonable uncertainty for the measurements.

Having convinced yourself that the results so far are reasonable, you can quickly finish the calculations. The best estimate for the length is the mean \bar{l} and the uncertainty is the SDOM $\sigma_{\bar{l}}$; so the final value for l is

$$l = 24.245 \pm 0.006 \text{ mm (or } 0.025\%);$$

the number in parenthesis is the percentage uncertainty. Similarly, the value for b is

$$b = 50.368 \pm 0.008 \text{ mm (or } 0.016\%).$$

Finally, the best estimate for the area $A = lb$ is the product of these values, with a fractional uncertainty given by the quadratic sum of those in l and b (assuming the

errors are independent):

$$\begin{aligned}
 A &= (24.245 \text{ mm} \pm 0.025\%) \times (50.368 \text{ mm} \pm 0.016\%) \\
 &= 1221.17 \text{ mm}^2 \pm 0.03\% \\
 &= 1221.2 \pm 0.4 \text{ mm}^2.
 \end{aligned} \tag{4.18}$$

To arrive at the answer (4.18) for A , we calculated the averages \bar{l} and \bar{b} , each with an uncertainty equal to the standard deviation of its mean. We then calculated the area A as the product of \bar{l} and \bar{b} and found the uncertainty by propagation of errors. We could have proceeded differently. For instance, we could have multiplied the first measured value of l by the first value of b to give a first answer for A . Continuing in this way we could have calculated 10 answers for A and then have subjected these 10 answers to statistical analysis, calculating \bar{A} , σ_A , and finally $\sigma_{\bar{A}}$. If, however, the errors in l and b are independent and random, and if we make enough measurements, this alternative procedure will produce the same result as the first one.⁵

Example: Another Spring

As a second example, consider a case in which a statistical analysis cannot be applied to the direct measurements but can to the final answers. Suppose we wish to measure the spring constant k of a spring by timing the oscillations of a mass m fixed to its end. We know from elementary mechanics that the period for such oscillations is $T = 2\pi\sqrt{m/k}$. Thus, by measuring T and m , we can find k as

$$k = 4\pi^2 m/T^2. \tag{4.19}$$

The simplest way to find k is to take a single, accurately known mass m and make several careful measurements of T . For various reasons, however, timing T for several *different* masses m may be more interesting. (For example, in this way, we could check that $T \propto \sqrt{m}$ as well as measure k .) We might then get a set of readings such as those in the first two lines of Table 4.4.

Table 4.4. Measurement of spring constant k .

Mass m (kg)	0.513	0.581	0.634	0.691	0.752	0.834	0.901	0.950
Period T (s)	1.24	1.33	1.36	1.44	1.50	1.59	1.65	1.69
$k = 4\pi^2 m/T^2$	13.17	12.97	etc.					

It obviously makes no sense to average the various different masses in the top line (or the times in the second line) because they are *not* different measurements of the same quantity. Nor can we learn anything about the uncertainty in our measurements by comparing the different values of m . On the other hand, we can com-

⁵The second procedure has a certain illogic because there is no particular reason to associate the first measurement of l with the first measurement of b . Indeed, we might have measured l eight times and b twelve times; then we couldn't pair off values. Thus, our first procedure is logically preferable.

bine each value of m with its corresponding period T and calculate k , as in the final line of Table 4.4. Our answers for k in the bottom line *are* all measurements of the same quantity and so can be subjected to statistical analysis. In particular, our best estimate for k is the mean, $\bar{k} = 13.16 \text{ N/m}$, and our uncertainty is the standard deviation of the mean, $\sigma_{\bar{k}} = 0.06 \text{ N/m}$ (see Problem 4.20). Thus, the final answer, based on the data of Table 4.4, is

$$\text{spring constant } k = 13.16 \pm 0.06 \text{ N/m.} \quad (4.20)$$

If we had formed reasonable estimates of the uncertainties in our original measurements of m and T , we could also have estimated the uncertainty in k by using error propagation, starting from these estimates for δm and δT . In this case, it would be a good idea to compare the final uncertainties in k obtained by the two methods.

4.6 Systematic Errors

In the past few sections, I have been taking for granted that all systematic errors were reduced to a negligible level before serious measurements began. Here, I take up again the disagreeable possibility of appreciable systematic errors. In the example just discussed, we may have been measuring m with a balance that read consistently high or low, or our timer may have been running consistently fast or slow. Neither of these systematic errors will show up in the comparison of our various answers for the spring constant k . As a result, the standard deviation of the mean $\sigma_{\bar{k}}$ can be regarded as the *random component* δk_{ran} of the uncertainty δk but is certainly not the total uncertainty δk . Our problem is to decide how to estimate the *systematic component* δk_{sys} and then how to combine δk_{ran} and δk_{sys} to give the complete uncertainty δk .

No simple theory tells us what to do about systematic errors. In fact, the only theory of systematic errors is that they must be identified and reduced until they are much less than the required precision. In a teaching laboratory, however, this goal is often not attainable. Students often cannot check a meter against a better one to correct it, much less buy a new meter to replace an inadequate one. For this reason, some teaching laboratories establish a rule that, in the absence of more specific information, meters should be considered to have some definite systematic uncertainty. For example, the decision might be that all stopwatches have up to 0.5% systematic uncertainty, all balances up to 1%, all voltmeters and ammeters up to 3%, and so on.

Given rules of this kind, there are various possible ways to proceed. None can really be rigorously justified, and we describe just one approach here. (Problems 4.23 to 4.28 contain more examples.) In the last example in Section 4.5, the spring constant $k = 4\pi^2 m/T^2$ was found by measuring a series of values of m and the corresponding values of T . As we have seen, a statistical analysis of the various answers for k gives the random component of δk as

$$\delta k_{\text{ran}} = \sigma_{\bar{k}} = 0.06 \text{ N/m.} \quad (4.21)$$

Suppose now we have been told that the balance used to measure m and the clock used for T have systematic uncertainties up to 1% and 0.5%, respectively. We can

then find the systematic component of δk by propagation of errors; the only question is whether to combine the errors in quadrature or directly. Because the errors in m and T are surely independent and some cancellation is therefore possible, using the quadratic sum is probably reasonable⁶; this choice gives

$$\frac{\delta k_{\text{sys}}}{k} = \sqrt{\left(\frac{\delta m_{\text{sys}}}{m}\right)^2 + \left(2\frac{\delta T_{\text{sys}}}{T}\right)^2} \quad (4.22)$$

$$= \sqrt{(1\%)^2 + (1\%)^2} = 1.4\% \quad (4.23)$$

and hence

$$\begin{aligned} \delta k_{\text{sys}} &= k_{\text{best}} \times (1.4\%) \\ &= (13.16 \text{ N/m}) \times 0.014 = 0.18 \text{ N/m}. \end{aligned} \quad (4.24)$$

Now that we have estimates for both the random and systematic uncertainties in k , we must decide how to state our final conclusion for the spring constant k with its overall uncertainty. Because the method for combining δk_{ran} and δk_{sys} is not completely clear, many scientists leave the two components separate and state a final answer in the form

$$\begin{aligned} (\text{measured value of } k) &= k_{\text{best}} \pm \delta k_{\text{ran}} \pm \delta k_{\text{sys}} \\ &= 13.16 \pm 0.06 \pm 0.18 \text{ N/m} \end{aligned} \quad (4.25)$$

(all of which should probably be rounded to one decimal place). Alternatively, a case can be made that δk_{ran} and δk_{sys} should be combined in quadrature, in which case we could state a single, total uncertainty

$$\begin{aligned} \delta k &= \sqrt{(\delta k_{\text{ran}})^2 + (\delta k_{\text{sys}})^2} \\ &= \sqrt{(0.06)^2 + (0.18)^2} = 0.19 \text{ N/m} \end{aligned} \quad (4.26)$$

and replace the conclusion (4.25) by

$$\begin{aligned} (\text{measured value of } k) &= k_{\text{best}} \pm \delta k \\ &= 13.16 \pm 0.19 \text{ N/m} \end{aligned}$$

or, probably better, $13.2 \pm 0.2 \text{ N/m}$.

The expression (4.26) for δk cannot really be rigorously justified. Nor is the significance of the answer clear; for example, we probably cannot claim 68% confidence that the true answer lies in the range $\bar{k} \pm \delta k$. Nonetheless, the expression does at least provide a reasonable estimate of our total uncertainty, given that our apparatus has systematic uncertainties we could not eliminate. In particular, there is one important respect in which the answer (4.26) is realistic and instructive. We saw in Section 4.4 that the standard deviation of the mean $\sigma_{\bar{k}}$ approaches zero as the number of measurements N is increased. This result suggested that, if you have the

⁶Whether we should use the quadratic or ordinary sum really depends on what is meant by the statement that the balance has “up to 1% systematic uncertainty.” If it means the error is *certainly* no more than 1% (and likewise for the clock), then direct addition is appropriate, and δk_{sys} is then *certainly* no more than 2%. On the other hand, perhaps an analysis of all balances in the laboratory has shown that they follow a normal distribution, with 68% of them better than 1% reliable (and likewise for the clocks). In this case, we can use addition in quadrature as in (4.22) with the usual significance of 68% confidence.

patience to make an enormous number of measurements, you can reduce the uncertainties indefinitely without having to improve your equipment or technique. We can now see that this suggestion is incorrect. Increasing N can reduce the *random* component $\delta k_{\text{ran}} = \sigma_{\bar{k}}$ indefinitely. But any given apparatus has *some* systematic uncertainty, which is *not* reduced as we increase N . From (4.26) we clearly see that little is gained from further reduction of δk_{ran} , once δk_{ran} is smaller than δk_{sys} . In particular, the total δk can never be made less than δk_{sys} . This fact simply confirms what we already guessed, that in practice a large reduction of the uncertainty requires improvements in techniques or equipment to reduce both the random and systematic errors in each single measurement.

As discussed in Chapter 2, a peculiar feature of the teaching laboratory is that you will probably be asked to measure quantities, such as the acceleration of gravity, for which an accurate, accepted value is already known. In this kind of experiment, the logic of the error analysis is a bit confusing. Probably the most honest course is to ignore the known accepted value until *after* you have done all calculations of your measured value, q_{best} , and its uncertainty. Then, of course, you must ask whether the accepted value lies inside (or at least close to) the range $q_{\text{best}} \pm \delta q$. If it does, you can simply record this agreement in your report. If the accepted value lies well outside the range $q_{\text{best}} \pm \delta q$, however, you have to examine the possible causes of the excessive discrepancy. For example, you might measure g , the acceleration of gravity, and get the results (all in m/s²),

$$g_{\text{best}} = 9.97, \quad (4.27)$$

with uncertainties

$$\delta g_{\text{ran}} = 0.02 \quad \text{and} \quad \delta g_{\text{sys}} = 0.03,$$

and hence a total uncertainty, as in (4.26), of

$$\delta g = 0.04.$$

Clearly, the accepted value of

$$g = 9.80 \text{ m/s}^2$$

lies far outside the measured range, 9.97 ± 0.04 . (More specifically, the discrepancy is 0.17, which is four times the uncertainty.) This result is definitely *not* satisfactory and further analysis is required.

The first thing to check is the possibility that you made a downright mistake in calculating g_{best} or one of the uncertainties δg_{ran} and δg_{sys} . If you can convince yourself that all your calculations were correct, the next possibility is that the accepted value is wrong. In the case of $g = 9.80 \text{ m/s}^2$ this possibility is rather unlikely, but it is entirely possible for plenty of other cases. For example, suppose you were measuring the density of air; because this is strongly dependent on the temperature and pressure, you could easily have looked up the wrong accepted value for this parameter.

Once you have eliminated these suspects, only one possibility is left: You must have overlooked some systematic error so that your value of δg_{sys} is too small. Ideally, you should try to find the culprit, but this search can be hard because of the many possibilities:

(1) Perhaps one of your meters had larger systematic errors than you had allowed for when you calculated δg_{sys} . You can investigate this possibility by determining how large a systematic error in your clock (or voltmeter, or whatever) would be needed to account for the offending discrepancy. If the needed error is not unreasonably large, you have one possible explanation of your difficulty.

(2) Another possible cause of systematic error is that you used an incorrect value for some parameter needed in your calculations. A celebrated example of this was Millikan's famous measurement of the electron's charge, e . Millikan's method depended on the viscosity of air, for which he used a value that was 0.4% too small. This discrepancy caused all of his values of e to be 0.6% too small, an error that was not noticed for nearly 20 years. This kind of mistake sometimes arises in a teaching laboratory when a student uses a value that has too few significant figures. For example, suppose you do an experiment with protons and you expect to have an accuracy better than 1%. If you take the proton's mass to be 1.7×10^{-27} kg (instead of the more exact 1.67×10^{-27} kg), you will have introduced a 2% systematic error, which will almost certainly frustrate your hope for 1% results.

(3) Much harder to analyze is the possibility of a flaw in the design of the experiment. For example, if you had measured g by dropping an object from a great height, air resistance could introduce an appreciable systematic error. [Note, however, that this error would not account for the large value of g in (4.27) because air resistance would cause an acceleration that was too *small*.] Similarly, if you try to measure the half-life of a radioactive material and your sample is contaminated with another material of shorter half-life, you will get an answer that is systematically too short.

Obviously, tracking down the source of systematic errors is difficult and has defied the best efforts of many great scientists. In all probability, your instructors are not going to penalize you too severely if you fail to do so. Nevertheless, they will expect an intelligent discussion of the problem and at least an honest admission that there appear to have been systematic errors that you were unable to identify.

Principal Definitions and Equations of Chapter 4

Suppose that we make N measurements, x_1, x_2, \dots, x_N of the same quantity x , all using the same method. Provided all uncertainties are random and small, we have the following results:

THE MEAN

The best estimate for x , based on these measurements, is their mean:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i. \quad [\text{See (4.5)}]$$

THE STANDARD DEVIATION

The average uncertainty of the individual measurements x_1, x_2, \dots, x_N is given by the standard deviation, or SD:

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum (x_i - \bar{x})^2}. \quad [\text{See (4.9)}]$$

This definition of the SD, often called the *sample* standard deviation, is the most appropriate for our purposes. The *population* standard deviation is obtained by replacing the factor $(N - 1)$ in the denominator by N . You will usually want to calculate standard deviations using the built-in function on your calculator; be sure you know which definition it uses.

The detailed significance of the standard deviation σ_x is that approximately 68% of the measurements of x (using the same method) should lie within a distance σ_x of the true value. (This claim is justified in Section 5.4.) This result is what allows us to identify σ_x as the *uncertainty* in any one measurement of x ,

$$\delta x = \sigma_x,$$

and, with this choice, we can be 68% confident that any one measurement will fall within σ_x of the correct answer.

THE STANDARD DEVIATION OF THE MEAN

As long as systematic uncertainties are negligible, the uncertainty in our best estimate for x (namely \bar{x}) is the standard deviation of the mean, or SDOM,

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}}. \quad [\text{See (4.14)}]$$

If there *are* appreciable systematic errors, then $\sigma_{\bar{x}}$ gives the *random component* of the uncertainty in our best estimate for x :

$$\delta x_{\text{ran}} = \sigma_{\bar{x}}.$$

If you have some way to estimate the systematic component δx_{sys} , a reasonable (but not rigorously justified) expression for the total uncertainty is the quadratic sum of δx_{ran} and δx_{sys} :

$$\delta x_{\text{tot}} = \sqrt{(\delta x_{\text{ran}})^2 + (\delta x_{\text{sys}})^2}. \quad [\text{See (4.26)}]$$

Problems for Chapter 4

For Section 4.2: The Mean and Standard Deviation

- 4.1. ★** You measure the time for a ball to drop from a second-floor window three times and get the results (in tenths of a second):

11, 13, 12.

(a) Find the mean and the standard deviation. For the latter, use both the “improved” definition (4.9) (the *sample* standard deviation) and the original definition (4.6) (the *population* standard deviation). Note how, even with only three measurements, the difference between the two definitions is not very big. (b) If you don’t yet know how to calculate the mean and standard deviation using your calculator’s built-in functions, take a few minutes to learn. Use your calculator to check your answers to part (a); in particular, find out which definition of the standard deviation your calculator uses.

4.2. ★ A student measures g , the acceleration of gravity, five times, with the results (all in m/s^2):

$$9.9, 9.6, 9.5, 9.7, 9.8.$$

(a) Find her mean and the standard deviation [as defined by (4.9)]. Do the calculation yourself using a layout similar to that in Table 4.2. (b) Check your value for the SD using the built-in function on your calculator. (If you don’t yet know how to calculate the mean and standard deviation using your calculator’s built-in functions, take a few minutes to learn.)

4.3. ★ Find the mean and standard deviation of the 10 measurements reported in (4.10). If you haven’t done either of the two previous problems, be sure to do this calculation yourself, then check it using the built-in functions on your calculator.

4.4. ★ The mean of N quantities x_1, \dots, x_N is defined as their sum divided by N ; that is, $\bar{x} = \sum x_i/N$. The deviation of x_i is the difference $d_i = x_i - \bar{x}$. Show clearly that the mean of the deviations d_1, \dots, d_N is always zero.

If you are not used to the Σ notation, you might want to do this problem both without and with the notation. For example, write out the sum $\sum(x_i - \bar{x})$ as $(x_1 - \bar{x}) + (x_2 - \bar{x}) + \dots + (x_N - \bar{x})$ and regroup the terms.

4.5. ★★ (a) Computing the standard deviation σ_x of N measurements x_1, \dots, x_N of a single quantity x requires that you compute the sum $\sum(x_i - \bar{x})^2$. Prove that this sum can be rewritten as

$$\sum[(x_i - \bar{x})^2] = \sum[(x_i)^2] - \frac{1}{N}[\sum(x_i)]^2. \quad (4.28)$$

This problem is a good exercise in using the Σ notation. Many calculators use the result to compute the standard deviation for the following reason: To use the expression on the left, a calculator must keep track of all the data (which uses a lot of memory) to calculate \bar{x} and then the sum indicated; to use the expression on the right, the machine needs only to keep a running total of $\sum(x_i^2)$ and $\sum(x_i)$, which uses much less memory. (b) Verify the identity (4.28) for the three measurements of Problem 4.1.

4.6. ★★ In Chapter 3, you learned that in a counting experiment, the uncertainty associated with a counted number is given by the “square-root rule” as the square root of that number. This rule can now be made more precise with the following statements (proved in Chapter 11): If we make several counts

$$\nu_1, \nu_2, \dots, \nu_N$$

of the number ν of random events that occur in a time T , then: (1) the best estimate for the true average number that occur in time T is the mean $\bar{\nu} = \sum \nu_i/N$ of our measurements, and (2) the *standard deviation* of the observed numbers should be approximately equal to the *square root* of this same best estimate; that is, the uncertainty in each measurement is $\sqrt{\bar{\nu}}$. In particular, if we make only one count ν , the best estimate is just ν and the uncertainty is the square root $\sqrt{\nu}$; this result is just the square-root rule of Chapter 3 with the additional information that the “uncertainty” is actually the standard deviation and gives the margins within which we can be approximately 68% confident the true answer lies. This problem and Problem 4.7 explore these ideas.

A nuclear physicist uses a Geiger counter to monitor the number of cosmic-ray particles arriving in his laboratory in any two-second interval. He counts this number 20 times with the following results:

$$\begin{array}{ccccccccccccc} 10, & 13, & 8, & 15, & 8, & 13, & 14, & 13, & 19, & 8, \\ 13, & 13, & 7, & 8, & 6, & 8, & 11, & 12, & 8, & 7. \end{array}$$

(a) Find the mean and standard deviation of these numbers. (b) The latter should be approximately equal to the square root of the former. How well is this expectation borne out?

4.7. ★★★ Read the first paragraph of Problem 4.6 and then do the following problem: A health physicist is testing a new detector and places it near a weak radioactive sample. In five separate 10-second intervals, the detector counts the following numbers of radioactive emissions:

$$16, 21, 13, 12, 15.$$

(a) Find the mean and standard deviation of these five numbers. (b) Compare the standard deviation with its expected value, the square root of the average number. (c) Naturally, the two numbers in part (b) do not agree exactly, and we would like to have some way to assess their disagreement. This problem is, in fact, one of error propagation. We have measured the number ν . The expected standard deviation in this number is just $\sqrt{\nu}$, a simple function of ν . Thus, the uncertainty in the standard deviation can be found by error propagation. Show, in this way, that the uncertainty in the SD is 0.5. Do the numbers in part (b) agree within this uncertainty?

4.8. ★★★ Spreadsheet programs, such as Lotus 123 or Excel, provide an excellent way to find the standard deviation of a set of measurements (and to record and process many kinds of data in general). (a) If you have access to one of these programs, create a spreadsheet to calculate the SD of any set of 10 measurements using the layout of Table 4.2. The first column should list the trial number i , and the second column should be where the user will enter the data x_i . The mean \bar{x} will be calculated by a formula that gives

$$(\text{sum of data entries})/(\text{number of data entries}),$$

which you can place to the side of the main table. In the third column, put a formula to calculate the deviation, $x_i - \bar{x}$, and in the fourth, a formula to calculate the deviation squared. Finally, somewhere to the side of the main table, write a formula to

find the SD as

$$\text{SD} = \sqrt{\frac{\text{sum of squared deviations}}{N - 1}}$$

(Most spreadsheet programs can calculate standard deviations automatically, but the point here is for you to create your own program to do it, not to use the built-in functions.) (b) Test your spreadsheet on the data of Equation (4.10). (c) (The hard part.) Most simple solutions to part (a) have two irritating drawbacks. First, as you enter each of the data, the spreadsheet will calculate answers for the mean and SD, but these answers will probably be incorrect (even for the data entered so far) until you have entered *all* the data. Second, your spreadsheet will probably need some modification before you can use it to find the mean and SD of a different number of data. If you can, modify your spreadsheet so that it does not suffer these defects and can find the mean and SD of *any number* of data (up to some convenient maximum, say 30). (d) Test your new spreadsheet using the data of Problem 4.6, and convince yourself that as you enter the data the program gives the correct answer at each stage.

For Section 4.3: The Standard Deviation as the Uncertainty in a Single Measurement

4.9. ★ A student measures the period of a pendulum three times and gets the answers 1.6, 1.8, and 1.7, all in seconds. What are the mean and standard deviation? [Use the improved definition (4.9) of the standard deviation.] If the student decides to make a fourth measurement, what is the probability that this new measurement will fall outside the range of 1.6 to 1.8 s? (The numbers here were chosen to “come out right.” In Chapter 5, I will explain how to do this kind of problem even when the numbers don’t come out right.)

4.10. ★ After several measurements of a quantity x , we expect to find that approximately 68% of the measurements fall within the range of $\bar{x} \pm \sigma_x$. (The exact number is 68.27 . . . %, but the difference is usually insignificant.) If you have not already done so, check the mean and standard deviation of the 10 measurements in (4.10). How many of the measurements should we expect to fall within the range of $\bar{x} \pm \sigma_x$? How many do?

4.11. ★ A student has to measure several unknown charges Q on a capacitor by discharging it through a ballistic galvanometer. The discharge kicks the galvanometer’s needle, whose resulting maximum swing tells the student the value of Q . Before making all her measurements, the student wants to know how reliably she can read the maximum displacement of the swinging needle. Therefore, she arranges to charge the capacitor to exactly the same voltage (and hence the same charge) five times and to measure the resulting charge. Her results (in microcoulombs) are:

$$1.2, 1.4, 1.6, 1.6, 1.2.$$

Based on these results, what would you suggest she take for the uncertainty δQ in her subsequent measurements of the charge Q , assuming she wants to be 68% confident that the correct value is within $\pm \delta Q$ of her measurement?

4.12. ★ To calibrate a prism spectrometer, a student sends light of 10 different known wavelengths λ through the spectrometer and measures the angle θ by which each beam is deflected. Using these results, he makes a calibration curve like Figure 3.8 showing λ as a function of θ . For the first value of λ , he measures θ six times and obtains these results (in degrees):

$$52.5, 52.3, 52.6, 52.5, 52.7, 52.4.$$

For each of the nine remaining values of λ , he measures the corresponding value of θ just once. What should he take for the uncertainty in each of these nine measurements of θ ?

4.13. ★★ (a) Calculate the mean and standard deviation for the following 30 measurements of a time t (in seconds):

$$\begin{aligned} & 8.16, 8.14, 8.12, 8.16, 8.18, 8.10, 8.18, 8.18, 8.18, 8.24, \\ & 8.16, 8.14, 8.17, 8.18, 8.21, 8.12, 8.12, 8.17, 8.06, 8.10, \\ & 8.12, 8.10, 8.14, 8.09, 8.16, 8.16, 8.21, 8.14, 8.16, 8.13. \end{aligned}$$

(You should certainly use the built-in functions on your calculator (or the spreadsheet you created in Problem 4.8 if you did), and you can save some button pushing if you drop all the leading 8s and shift the decimal point two places to the right before doing any calculation.) (b) We know that after several measurements, we can expect about 68% of the observed values to be within σ_t of \bar{t} (that is, inside the range $\bar{t} \pm \sigma_t$). For the measurements of part (a), about how many would you expect to lie *outside* the range $\bar{t} \pm \sigma_t$? How many do? (c) In Chapter 5, I will show that we can also expect about 95% of the values to be within $2\sigma_t$ of \bar{t} (that is, inside the range $\bar{t} \pm 2\sigma_t$). For the measurements of part (a), about how many would you expect to lie *outside* the range $\bar{t} \pm 2\sigma_t$? How many do?

4.14. ★★ (a) If you have not yet done it, do Problem 4.6. (b) About how many of the measurements of this problem should you expect to lie outside the range $\bar{v} \pm \sigma_v$? How many do? (c) In Chapter 5, I will show that we can expect about 95% of the measurements to be within $2\sigma_v$ of \bar{v} (that is, inside the range $\bar{v} \pm 2\sigma_v$). For the measurements of part (a), about how many would you expect to lie *outside* the range $\bar{v} \pm 2\sigma_v$? How many do?

For Section 4.4: The Standard Deviation of the Mean

4.15. ★ Given the three measurements in Problem 4.1, what should you state for your best estimate for the time concerned and its uncertainty? (Your answer will illustrate how the mean can have more significant figures than the original measurements.)

4.16. ★ (a) Based on the five measurements of g reported in Problem 4.2, what should be the student's best estimate for g and its uncertainty? (b) How well does her result agree with the accepted value of 9.8 m/s^2 ?

4.17. ★ (a) Based on the 30 measurements in Problem 4.13, what would be your best estimate for the time involved and its uncertainty, assuming all uncertainties

are random? (b) Comment on the number of significant digits in your best estimate, as compared with the number of significant digits in the data.

4.18. ★ After measuring the speed of sound u several times, a student concludes that the standard deviation σ_u of her measurements is $\sigma_u = 10$ m/s. If all uncertainties were truly random, she could get any desired precision by making enough measurements and averaging. (a) How many measurements are needed to give a final uncertainty of ± 3 m/s? (b) How many for a final uncertainty of only ± 0.5 m/s?

4.19. ★★★ (a) The data in Problem 4.6 are 20 measurements of the number of cosmic-ray particles counted by a Geiger counter in 2 seconds. By averaging these numbers, find the best estimate for the number of particles that arrive in 2 seconds and the uncertainty in that number (as given by the SDOM). (b) Another way to do this problem is as follows: If you add all the data, you will have the number counted in 40 seconds (with an uncertainty given by the square root of that number). If you now divide this result by 20, you should get the number of particles that arrive in 2 seconds and its uncertainty. Check that your answers in (b) and (c) agree, at least approximately. (Assume that the 2-second time intervals are measured with negligible uncertainty.) (c) Prove that these two methods should, ideally, give the same answers. (Your proof should be general; that is, you should work in terms of algebraic symbols, not the specific numbers of Problem 4.6. Remember that the SD of any counted number should be the square root of that number.)

For Section 4.5: Examples

4.20. ★ Complete the calculations of the spring constant k in Table 4.4. Then compute \bar{k} and its uncertainty (the SDOM, $\sigma_{\bar{k}}$).

4.21. ★ Table 4.3 records 10 measurements each of the length l and breadth b of a rectangle. These values were used to calculate the area $A = lb$. If the measurements were made in pairs (one of l and one of b), it would be natural to multiply each pair together to give a value of A —the first l times the first b to give a first value of A , and so on. Calculate the resulting 10 values of A , the mean \bar{A} , the SD σ_A , and the SDOM $\sigma_{\bar{A}}$. Compare the answers for \bar{A} and $\sigma_{\bar{A}}$ with the answer (4.18) obtained by calculating the averages \bar{l} and \bar{b} and then taking A to be $\bar{l}\bar{b}$, with an uncertainty given by error propagation. (For a large number of measurements, the two methods should agree.)

4.22. ★★ This problem is an example of an experiment for which the error analysis can be done in either of two ways: by propagating the estimated errors in the original measurements or by doing a statistical analysis of the various answers. A student wants to measure g , the acceleration of gravity, using a simple pendulum, as described briefly in the introduction to this chapter. Because the period is known to be $T = 2\pi\sqrt{l/g}$, where l is the length of the pendulum, she can find g as $g = 4\pi^2 l/T^2$. She measures T for five different values of l and obtains the following results:

Length, l (cm):	57.3	61.1	73.2	83.7	95.0
Time, T (s):	1.521	1.567	1.718	1.835	1.952

(a) Copy the table of data and add a row in which you list her five computed values of g . (b) She estimates she can read the lengths l within about 0.3% (that is, two or three millimeters).⁷ Similarly, she estimates that all of the times are within $\pm 0.2\%$. Use error propagation to find the uncertainty in her values for g . (c) Because her values of g are five measurements of the same quantity, we can analyze them statistically. In particular, their standard deviation should represent the uncertainty in any one of her answers. What is the SD, and how does it compare with the uncertainty found by error propagation in part (b)? [You should not expect the agreement to be *especially* good because we don't know the exact nature of her original estimated uncertainties (nor, probably, does she). Nevertheless, the two methods should agree *roughly*, and a large disagreement would be a clear signal that something had gone wrong.] (d) What is her final answer for g with its uncertainty? [Use the statistical analysis of part (c), and remember that the final uncertainty is the SDOM. How does her answer compare with the accepted value (in her laboratory) of 979.6 cm/s^2 ?

For Section 4.6: Systematic Errors

4.23. ★ A famous example of a systematic error occurred in Millikan's historic measurement of the electron's charge e . He worked on this experiment for several years and had reduced all random errors to a very low level, certainly less than 0.1%. Unfortunately, his answer for e depended on the viscosity of air (denoted η), and the value of η that he used was 0.4% too low. His value for e had the form $e = K\eta^{3/2}$, where K stands for a complicated expression involving several measured parameters but not η . Therefore, the systematic error in η caused a systematic error in e . Given that all other errors (random and systematic) were much less than 0.4%, what was his error in e ? (This example is typical of many systematic errors. Until the errors are identified, nothing can be done about them. Once identified, the errors can be eliminated, in this case by using the right value of η .)

4.24. ★★ In some experiments, systematic errors can be caused by the neglect of an effect that is *not* (in the situation concerned) negligible, for example, neglect of heat losses from a badly insulated calorimeter or neglect of friction for a poorly lubricated cart. Here is another example: A student wants to measure the acceleration of gravity g by timing the fall of a wooden ball (3 or 4 inches across) dropped from four different windows in a tall building. He assumes that air resistance is negligible and that the distance fallen is given by $d = \frac{1}{2}gt^2$. Using a tape measure and an electric timer, he measures the distances and times of the four separate drops as follows:

Distance, d (meters):	15.43	17.37	19.62	21.68
Time, t (seconds):	1.804	1.915	2.043	2.149

(a) Copy these data and add a third row in which you put the corresponding accelerations, calculated as $g = 2d/t^2$. (b) Based on these results, what is his best estimate

⁷Although the percent uncertainties in the five measurements of l are probably not exactly the same, it is appropriate (and time saving) in many experiments to assume they are at least approximately so. In other words, instead of doing five separate error propagations, you may often appropriately do just one for a representative case and assume that all five cases are reasonably similar.

for g , assuming that all errors are random? Show that this answer is inconsistent with the accepted value of $g = 9.80 \text{ m/s}^2$. (c) Having checked his calculations, tape measure, and timer, he concludes (correctly) that there must be some systematic error causing an acceleration different from 9.80 m/s^2 , and he suggests that air resistance is probably the culprit. Give at least two arguments to support this suggestion. (d) Suggest a couple of ways he could modify the experiment to reduce the effect of this systematic error.

4.25. ★★ (a) A student measures the speed of sound as $u = f\lambda$, where f is the frequency shown on the dial of an audio oscillator, and λ is the wavelength measured by locating several maxima in a resonant air column. Because there are several measurements of λ , they can be analyzed statistically, and the student concludes that $\lambda = 11.2 \pm 0.5 \text{ cm}$. Only one measurement has been taken of $f = 3,000 \text{ Hz}$ (the setting on the oscillator), and the student has no way to judge its reliability. The instructor says that the oscillator is “certainly 1% reliable”; therefore, the student allows for a 1% systematic error in f (but none in λ). What is the student’s answer for u with its uncertainty? Is the possible 1% systematic error from the oscillator’s calibration important? (b) If the student’s measurement had been $\lambda = 11.2 \pm 0.1 \text{ cm}$ and the oscillator calibration had been 3% reliable, what would the answer have been? Is the systematic error important in this case?

4.26. ★★ A student wants to check the resistance of a resistor by measuring the voltage across it (V) and the resulting current through it (I) and then calculating the resistance as $R = V/I$. He measures four different values of V and the corresponding currents I , as follows:

Voltage, V (volts):	11.2	13.4	15.1	17.7
Current, I (amps):	4.67	5.46	6.28	7.22

(a) Calculate the four corresponding values of R (which will come out in ohms). What is his best estimate for R , and what is the random component of its uncertainty (δR_{ran})? (b) The resistor is rated at 2.50 ohms, which does not lie within the range $R_{\text{best}} \pm \delta R_{\text{ran}}$, so he considers the possibility that the voltmeter and ammeter suffer some systematic error. The laboratory technician states that many of the meters in the laboratory have up to 2% systematic error. Use error propagation to find the possible systematic error in R , and then combine the systematic and random errors to give the total uncertainty. (In both calculations, combine the errors in quadrature.) What is his final answer and how does it compare with the given value?

4.27. ★★ Some experiments require calibration of the equipment before the measurements can be made. Any random errors in the calibration will usually become systematic errors in the experiment itself. To illustrate this effect, consider an experiment on the Zeeman effect, in which a magnetic field causes a tiny shift in the frequency of light given out by an atom. This shift can be measured using a Fabry-Perot interferometer, which consists of two parallel reflecting surfaces a distance d apart (where d is typically a couple of millimeters). To use the interferometer, one must know the distance d ; that is, one must calibrate the instrument by measuring d . A convenient way to make this measurement is to send light of an accurately known wavelength λ through the interferometer, which produces a series of interfer-

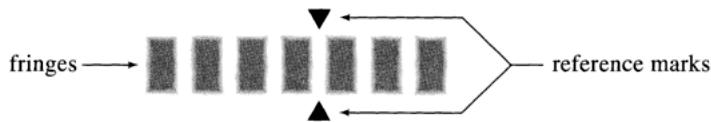


Figure 4.3. When light of one wavelength is sent through a Fabry-Perot interferometer, it produces a pattern of alternating light and dark fringes. If the air is pumped out of the interferometer, the whole pattern shifts sideways, and the number of complete fringes (light-dark-light) that pass the reference marks can be counted; for Problem 4.27.

ence fringes like those depicted in Figure 4.3. If all the air is then pumped out of the chamber that houses the interferometer, the interference pattern slowly shifts sideways, and the number of fringes that move past the reference marks is $N = 2(n - 1)d/\lambda$, where n denotes the refractive index of air. Because n and λ are known accurately, d can be found by counting N . Because N is not necessarily an integer, the fractions of complete fringes that pass the reference marks must be estimated, and this estimation introduces the only serious source of uncertainty.

(a) In one such experiment, a student measures N five times as follows:

$$\text{values of } N = 3.0, 3.5, 3.2, 3.0, 3.2.$$

What is her best estimate for N and its uncertainty? What is the resulting percent uncertainty in d ? (b) The uncertainty you just found is purely random. Nevertheless, in all subsequent measurements using the interferometer, she will be using the same value of d she found in part (a), and any error in that value will cause a *systematic* error in her final answers. What percent value should she use for this systematic error in d ?

4.28. ★★ Systematic errors sometimes arise when the experimenter unwittingly measures the wrong quantity. Here is an example: A student tries to measure g using a pendulum made of a steel ball suspended by a light string. (See Figure 4.4.) He

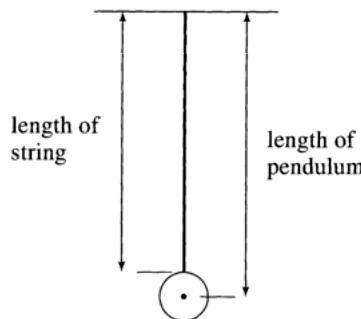


Figure 4.4. A pendulum consists of a metal ball suspended by a string. The effective length of the pendulum is the length of the string *plus* the radius of the ball; for Problem 4.28.

records five different lengths of the pendulum l and the corresponding periods T as follows:

Length, l (cm):	51.2	59.7	68.2	79.7	88.3
Period, T (s):	1.448	1.566	1.669	1.804	1.896

- (a) For each pair, he calculates g as $g = 4\pi^2 l/T^2$. He then calculates the mean of these five values, their SD, and their SDOM. Assuming all his errors are random, he takes the SDOM as his final uncertainty and quotes his answer in the standard form of mean \pm SDOM. What is his answer for g ? (b) He now compares his answer with the accepted value $g = 979.6 \text{ cm/s}^2$ and is horrified to realize that his discrepancy is nearly 10 times larger than his uncertainty. Confirm this sad conclusion. (c) Having checked all his calculations, he concludes that he must have overlooked some systematic error. He is sure there was no problem with the measurement of the period T , so he asks himself the question: How large would a systematic error in the length l have to be so that the margins of the total error *just included* the accepted value 979.6 cm/s^2 ? Show that the answer is approximately 1.5%. (d) This result would mean that his length measurements suffered a systematic error of about a centimeter—a conclusion he first rejects as absurd. As he stares at the pendulum, however, he realizes that 1 cm is about the radius of the ball and that the lengths he recorded were the lengths *of the string*. Because the correct length of the pendulum is the distance from the pivot to the *center* of the ball (see Figure 4.4), his measurements were indeed systematically off by the radius of the ball. He therefore uses callipers to find the ball's diameter, which turns out to be 2.00 cm. Make the necessary corrections to his data and compute his final answer for g with its uncertainty.

Chapter 5

The Normal Distribution

This chapter continues our discussion of the statistical analysis of repeated measurements. Chapter 4 introduced the important ideas of the mean, the standard deviation, and the standard deviation of the mean; we saw their significance and some of their uses. This chapter supplies the theoretical justification for these statistical ideas and gives proofs of several results stated without proof in earlier chapters.

The first problem in discussing measurements repeated many times is to find a way to handle and display the values obtained. One convenient method is to use a *distribution* or *histogram*, as described in Section 5.1. Section 5.2 introduces the notion of the *limiting distribution*, the distribution of results that would be obtained if the number of measurements become infinitely large. In Section 5.3, I define the *normal distribution*, or *Gauss distribution*, which is the limiting distribution of results for any measurement subject to many small random errors.

Once the mathematical properties of the normal distribution are understood, we can proceed to prove several important results quite easily. Section 5.4 provides proof that, as anticipated in Chapter 4, about 68% of all measurements (all of one quantity and all using the same technique) should lie within one standard deviation of the true value. Section 5.5 proves the result, used back in Chapter 1, that if we make N measurements x_1, x_2, \dots, x_N of some quantity x , then our best estimate x_{best} based on these values is the mean $\bar{x} = \sum x_i / N$. Section 5.6 justifies the use of addition in quadrature when propagating errors that are independent and random. In Section 5.7, I prove that the uncertainty of the mean \bar{x} , when used as the best estimate of x , is given by the standard deviation of the mean $\sigma_{\bar{x}} = \sigma_x / \sqrt{N}$, as stated in Chapter 4. Finally, Section 5.8 discusses how to assign a numerical confidence to experimental results.

The mathematics used in this chapter is more advanced than used thus far. In particular, you will need to understand the basic ideas of integration—the integral as the area under a graph, changes of variables, and (occasionally) integration by parts. However, once you have worked through Section 5.3 on the normal distribution (going over calculations with a pencil and paper, if necessary) you should be able to follow the rest of the chapter without much difficulty.

5.1 Histograms and Distributions

It should be clear that the serious statistical analysis of an experiment requires us to make many measurements. Thus, we first need to devise methods for recording and displaying large numbers of measured values. Suppose, for instance, we were to make 10 measurements of some length x . For example, x might be the distance from a lens to an image formed by the lens. We might obtain the values (all in cm)

$$26, 24, 26, 28, 23, 24, 25, 24, 26, 25. \quad (5.1)$$

Written this way, these 10 numbers convey fairly little information, and if we were to record many more measurements this way, the result would be a confusing jungle of numbers. Obviously, a better system is needed.

As a first step, we can reorganize the numbers (5.1) in ascending order,

$$23, 24, 24, 24, 25, 25, 26, 26, 26, 28. \quad (5.2)$$

Next, rather than recording the three readings 24, 24, 24, we can simply record that we obtained the value 24 three times; in other words, we can record the *different* values of x obtained, together with the *number* of times each value was found, as in Table 5.1.

Table 5.1. Measured lengths x and their numbers of occurrences.

Different values, x_k	23	24	25	26	27	28
Number of times found, n_k	1	3	2	3	0	1

Here, I have introduced the notation x_k ($k = 1, 2, \dots$) to denote the various different values found: $x_1 = 23$, $x_2 = 24$, $x_3 = 25$, and so on. And n_k ($k = 1, 2, \dots$) denotes the number of times the corresponding value x_k was found: $n_1 = 1$, $n_2 = 3$, and so on.

If we record measurements as in Table 5.1, we can rewrite the definition of the mean \bar{x} in what proves to be a more convenient way. From our old definition, we know that

$$\bar{x} = \frac{\sum_i x_i}{N} = \frac{23 + 24 + 24 + 24 + 25 + \dots + 28}{10}. \quad (5.3)$$

This equation is the same as

$$\bar{x} = \frac{23 + (24 \times 3) + (25 \times 2) + \dots + 28}{10}$$

or in general

$$\bar{x} = \frac{\sum_k x_k n_k}{N}. \quad (5.4)$$

In the original form (5.3), we sum over *all* the measurements made; in (5.4) we sum over all *different* values obtained, multiplying each value by the number of times it occurred. These two sums are obviously the same, but the form (5.4) proves more useful when we make many measurements. A sum like that in (5.4) is sometimes called a *weighted sum*; each value x_k is *weighted* by the number of times it occurred, n_k . For later reference, note that if we add up all the numbers n_k , we obtain the total number of measurements made, N . That is,

$$\sum_k n_k = N. \quad (5.5)$$

(For example, for Table 5.1 this equation asserts that the sum of the numbers in the bottom line is 10.)

Quick Check 5.1. In his first two years at college, Joe takes 20 courses (all with the same number of credits) and earns 7 As, 4 Bs, 7 Cs, and 2 Fs. For the purpose of computing a grade point average (GPA), each letter grade is assigned a numerical score in the usual way, as follows:

Letter grade:	F	D	C	B	A
Score, s_k :	$s_1 = 0$	$s_2 = 1$	$s_3 = 2$	$s_4 = 3$	$s_5 = 4$

Set up a table like Table 5.1 showing the different possible scores s_k and the number of times n_k they were obtained. Use Equation (5.4) to compute Joe's GPA, \bar{s} .

The ideas of the past two paragraphs can be rephrased in a way that is often more convenient. Instead of saying that the result $x = 24$ was obtained three times, we can say that $x = 24$ was obtained in 3/10 of all our measurements. In other words, instead of using n_k , the *number* of times the result x_k occurred, we introduce the fraction

$$F_k = \frac{n_k}{N}, \quad (5.6)$$

which is the *fraction* of our N measurements that gave the result x_k . The fractions F_k are said to specify the *distribution* of our results because they describe how our measurements were *distributed* among the different possible values.

In terms of the fractions F_k , we can rewrite the formula (5.4) for the mean \bar{x} in the compact form

$$\bar{x} = \sum_k x_k F_k. \quad (5.7)$$

That is, the mean \bar{x} is just the weighted sum of all the different values x_k obtained, with each x_k weighted by the fraction of times it occurred, F_k .

The result (5.5) implies that

$$\sum_k F_k = 1. \quad (5.8)$$

That is, if we add up the fractions F_k for all possible results x_k , we must get 1. Any set of numbers whose sum is 1 is said to be *normalized*, and the relation (5.8) is therefore called the *normalization condition*.

The distribution of our measurements can be displayed graphically in a *histogram*, as in Figure 5.1. This figure is just a plot of F_k against x_k , in which the

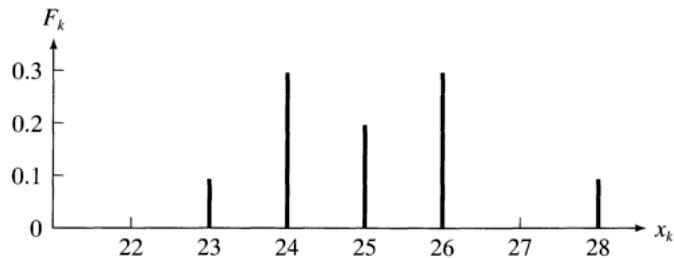


Figure 5.1. Histogram for 10 measurements of a length x . The vertical axis shows the fraction of times F_k that each value x_k was observed.

different measured values x_k are plotted along the horizontal axis and the fraction of times each x_k was obtained is indicated by the height of the vertical bar drawn above x_k . (We can also plot n_k against x_k , but for our purposes the plot of F_k against x_k is more convenient.) Data displayed in histograms like this one can be comprehended quickly and easily, as many writers for newspapers and magazines are aware.

A histogram like that in Figure 5.1 can be called a *bar histogram* because the distribution of results is indicated by the heights of the vertical bars above the x_k . This kind of histogram is appropriate whenever the values x_k are tidily spaced, with integer values. (For example, students' scores on an examination are usually integers and are displayed conveniently using a bar histogram.) Most measurements, however, do not provide tidy integer results because most physical quantities have a continuous range of possible values. For example, rather than the 10 lengths reported in Equation (5.1), you are much more likely to obtain 10 values like

$$26.4, 23.9, 25.1, 24.6, 22.7, 23.8, 25.1, 23.9, 25.3, 25.4. \quad (5.9)$$

A bar histogram of these 10 values would consist of 10 separate bars, all the same height, and would convey comparatively little information. Given measurements like those in (5.9), the best course is to divide the range of values into a convenient number of *intervals* or "bins," and to count how many values fall into each "bin." For example, we could count the number of the measurements (5.9) between $x = 22$ and 23, between $x = 23$ and 24, and so on. The results of counting in this way are

shown in Table 5.2. (If a measurement happens to fall exactly on the boundary between two bins, you must decide where to place it. A simple and reasonable course is to assign half a measurement to each of the two bins.)

Table 5.2. The 10 measurements (5.9) grouped in bins.

Bin Observations in bin	22 to 23	23 to 24	24 to 25	25 to 26	26 to 27	27 to 28
	1	3	1	4	1	0

The results in Table 5.2 can be plotted in a form we can call a *bin histogram*, as shown in Figure 5.2. In this plot, the fraction of measurements that fall in each bin is indicated by the area of the rectangle drawn above the bin. Thus, the shaded

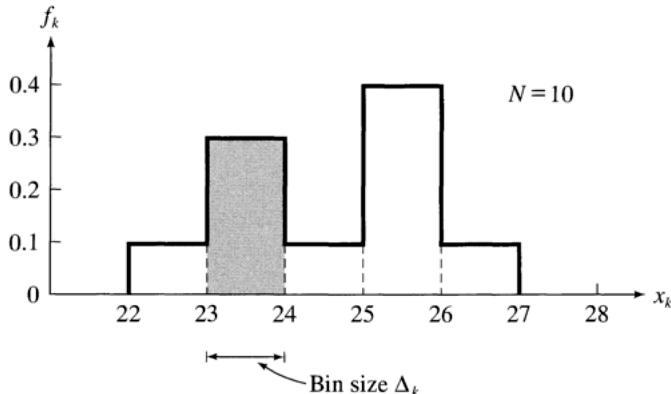


Figure 5.2. Bin histogram showing the fraction of the 10 measurements (5.9) of x that fall in the “bins” 22 to 23, 23 to 24, and so on. The area of the rectangle above each interval gives the fraction of measurements that fall in that interval. Thus, the area of the shaded rectangle is 0.3, indicating that $3/10$ of all measurements lie between 23 and 24.

rectangle above the interval from $x = 23$ to $x = 24$ has area $0.3 \times 1 = 0.3$, indicating that $3/10$ of all the measurements fell in this interval. In general, we denote the width of the k^{th} bin by Δ_k . (These widths are usually all the same, though they certainly don't have to be.) The height f_k of the rectangle drawn above this bin is chosen so that the area $f_k \Delta_k$ is

$$f_k \Delta_k = \text{fraction of measurements in } k^{\text{th}} \text{ bin.}$$

In other words, in a bin histogram the area $f_k \Delta_k$ of the k^{th} rectangle has the same significance as the height F_k of the k^{th} bar in a bar histogram.

Some care is needed in choosing the width Δ_k of the bins for a histogram. If the bins are made much too wide, then all the readings (or almost all) will fall in one bin, and the histogram will be an uninteresting single rectangle. If the bins are made too narrow, then few of them will contain more than one reading, and the histogram will consist of numerous narrow rectangles almost all of the same height.

Clearly, the bin width must be chosen so several readings fall in each of several bins. Thus, when the total number of measurements N is small, we have to choose our bins relatively wide, but if we increase N , then we can usually choose narrower bins.

Quick Check 5.2. A class of 20 students takes an exam, which is graded out of 50 points, and obtains the following results:

$$\begin{aligned} & 26, 33, 38, 41, 49, 28, 36, 38, 47, 41, \\ & 32, 37, 48, 44, 27, 32, 34, 44, 37, 30 \end{aligned}$$

(These scores were taken from an alphabetical list of the students.) On a piece of square-ruled paper, draw a bin histogram of the scores, using bin boundaries at 25, 30, 35, 40, 45, and 50. Label the vertical scale so that the area of each rectangle is the fraction of students in the corresponding bin.

5.2 Limiting Distributions

In most experiments, as the number of measurements increases, the histogram begins to take on a definite simple shape. This evolving shape is clearly visible in Figures 5.3 and 5.4, which show 100 and 1,000 measurements of the same quantity as in Figure 5.2. After 100 measurements, the histogram has become a single peak, which is approximately symmetrical. After 1,000 measurements, we have been able to halve the bin size, and the histogram has become quite smooth and regular. These three graphs illustrate an important property of most measurements. As the number of measurements approaches infinity, their distribution approaches some definite, continuous curve. When this happens, the continuous curve is called the *limiting distribution*.¹ Thus, for the measurements of Figures 5.2 through 5.4, the limiting

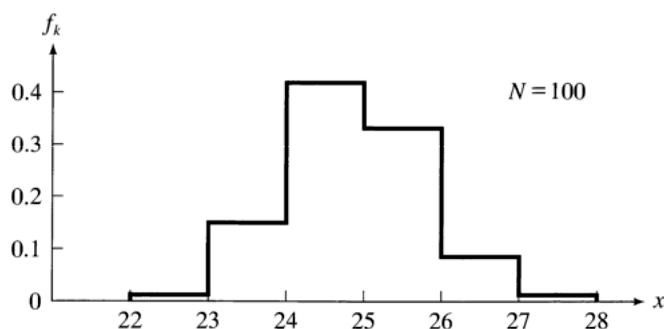


Figure 5.3. Histogram for 100 measurements of the same quantity as in Figure 5.2.

¹Some common synonyms (or approximate synonyms) for the limiting distribution are: parent distribution, infinite parent distribution, universe distribution, and parent population.

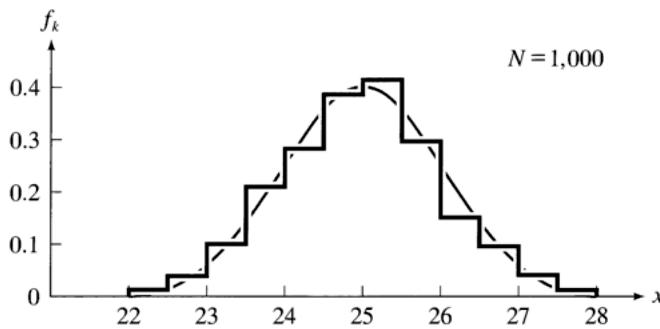


Figure 5.4. Histogram for 1,000 measurements of the same quantity as in Figure 5.3. The broken curve is the limiting distribution.

distribution appears to be close to the symmetric bell-shaped curve superimposed on Figure 5.4.

Note that the limiting distribution is a theoretical construct that can never itself be measured exactly. The more measurements we make, the closer our histogram approaches the limiting distribution. But only if we were to make an infinite number of measurements and use infinitesimally narrow bins would we actually obtain the limiting distribution itself. Nevertheless, there is good reason to believe that every measurement *does* have a limiting distribution to which our histogram approaches ever closer as we make more and more measurements.

A limiting distribution, such as the smooth curve in Figure 5.4, defines a function, which we call $f(x)$. The significance of this function is shown by Figure 5.5. As we make more and more measurements of the quantity x , our histogram will eventually be indistinguishable from the limiting curve $f(x)$. Therefore, the fraction of measurements that fall in any small interval x to $x + dx$ equals the area $f(x) dx$ of the shaded strip in Figure 5.5(a):

$$f(x) dx = \text{fraction of measurements that fall between } x \text{ and } x + dx. \quad (5.10)$$

More generally, the fraction of measurements that fall between any two values a and b is the total area under the graph between $x = a$ and $x = b$ (Figure 5.5b).

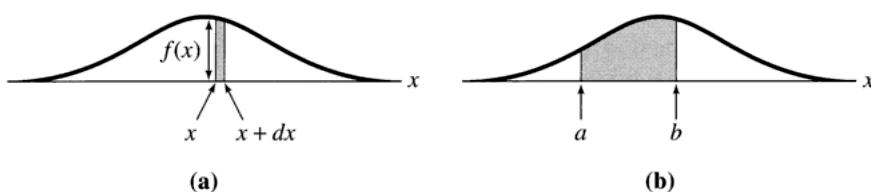


Figure 5.5. A limiting distribution $f(x)$. (a) After very many measurements, the fraction that falls between x and $x + dx$ is the area $f(x) dx$ of the narrow strip. (b) The fraction that falls between $x = a$ and $x = b$ is the shaded area.

This area is just the *definite integral* of $f(x)$. Thus, we have the important result that

$$\int_a^b f(x) dx = \text{fraction of measurements that fall between } x = a \text{ and } x = b. \quad (5.11)$$

Understanding the meaning of the two statements (5.10) and (5.11) is important. Both tell us the fraction of measurements expected to lie in some interval after we make a *very large number of measurements*. Another, very useful, way to say this is that $f(x) dx$ is the *probability* that a single measurement of x will give an answer between x and $x + dx$,

$$f(x) dx = \text{probability that any one measurement will give an answer between } x \text{ and } x + dx. \quad (5.12)$$

Similarly, the integral $\int_a^b f(x) dx$ tells us the probability that any one measurement will fall between $x = a$ and $x = b$. We have arrived at the following important conclusion: If we knew the limiting distribution $f(x)$ for the measurement of a given quantity x with a given apparatus, then we would know the probability of obtaining an answer in any interval $a \leq x \leq b$.

Because the total probability of obtaining an answer anywhere between $-\infty$ and $+\infty$ must be one, the limiting distribution $f(x)$ must satisfy

$$\int_{-\infty}^{\infty} f(x) dx = 1. \quad (5.13)$$

This identity is the natural analog of the normalization sum (5.8), $\sum_k F_k = 1$, and a function $f(x)$ satisfying (5.13) is said to be *normalized*.

The limits $\pm\infty$ in the integral (5.13) may seem puzzling. They do not mean that we really expect to obtain answers ranging all the way from $-\infty$ to ∞ . Quite the contrary. In a real experiment, the measurements all fall in some fairly small finite interval. For example, the measurements of Figure 5.4 all lie between $x = 21$ and $x = 29$. Even after infinitely many measurements, the fraction lying outside $x = 21$ to $x = 29$ would be entirely negligible. In other words, $f(x)$ is essentially zero outside this range, and it makes no difference whether the integral (5.13) runs from $-\infty$ to $+\infty$ or 21 to 29. Because we generally don't know what these finite limits are, for convenience we leave them as $\pm\infty$.

If the measurement under consideration is very precise, all the values obtained will be close to the actual value of x , so the histogram of results, and hence the limiting distribution, will be narrowly peaked like the solid curve in Figure 5.6. If the measurement is of low precision, then the values found will be widely spread and the distribution will be broad and low like the dashed curve in Figure 5.6.

The limiting distribution $f(x)$ for measurement of a given quantity x using a given apparatus describes how results would be distributed after many, many measurements. Thus, if we knew $f(x)$, we could calculate the mean value \bar{x} that would be found after many measurements. We saw in (5.7) that the mean of any number

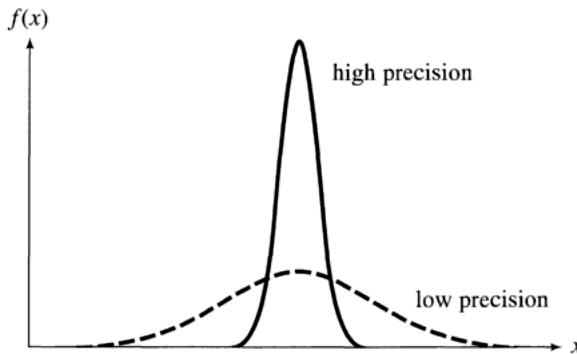


Figure 5.6. Two limiting distributions, one for a high-precision measurement, the other for a low-precision measurement.

of measurements is the sum of all different values x_k , each weighted by the fraction of times it is obtained:

$$\bar{x} = \sum_k x_k F_k. \quad (5.14)$$

In the present case, we have an enormous number of measurements with distribution $f(x)$. If we divide the whole range of values into small intervals x_k to $x_k + dx_k$, the fraction of values in each interval is $F_k = f(x_k) dx_k$ and in the limit that all intervals go to zero, (5.14) becomes

$$\bar{x} = \int_{-\infty}^{\infty} x f(x) dx. \quad (5.15)$$

Remember that this formula gives the mean \bar{x} expected after infinitely many trials.

Similarly, we can calculate the standard deviation σ_x obtained after many measurements. Because we are concerned with the limit $N \rightarrow \infty$, it makes no difference which definition of σ_x we use, the original (4.6) or the “improved” (4.9) with N replaced by $N - 1$. In either case, when $N \rightarrow \infty$, σ_x^2 is the average of the squared deviation $(x - \bar{x})^2$. Thus, exactly the argument leading to (5.15) gives, after many trials,

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x) dx \quad (5.16)$$

(see Problem 5.10).

5.3 The Normal Distribution

Different types of measurements have different limiting distributions. Not all limiting distributions have the symmetric bell shape illustrated in Section 5.2. (For

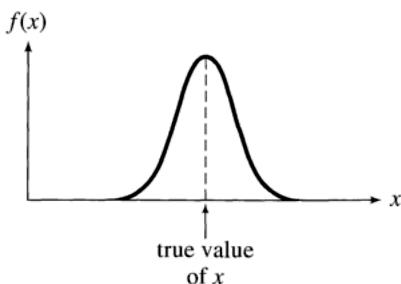


Figure 5.7. The limiting distribution for a measurement subject to many small random errors. The distribution is bell-shaped and centered on the true value of the measured quantity x .

example, the binomial and Poisson distributions discussed in Chapters 10 and 11 are usually not symmetric.) Nevertheless, many measurements are found to have a symmetric bell-shaped curve for their limiting distribution. In fact, I will prove in Chapter 10 that if a measurement is subject to many small sources of random error and negligible systematic error, the measured values will be distributed in accordance with a bell-shaped curve and this curve will be centered on the true value of x , as in Figure 5.7. In the remainder of this chapter, I will confine my attention to measurements with this property.

If our measurements have appreciable systematic errors, we would *not* expect the limiting distribution to be centered on the true value. Random errors are equally likely to push our readings above or below the true value. If all errors are random, after many measurements the number of observations above the true value will be the same as that below it, and our distribution of results will therefore be centered on the true value. But a systematic error (such as that caused by a tape measure that is stretched or a clock that runs slow) pushes all values in one direction and so pushes the distribution of observed values off center from the true value. In this chapter, I will assume that the distribution is centered on the true value. This is equivalent to assuming that all systematic errors have been reduced to a negligible level.

I now turn briefly to a question we have avoided discussing so far: What is the “true value” of a physical quantity? This question is a hard one that has no satisfactory, simple answer. Because no measurement can exactly determine the true value of any continuous variable (a length, a time, etc.), whether the true value of such a quantity exists is not even clear. Nevertheless, I will make the convenient assumption that every physical quantity does have a true value.

We can think of the true value of a quantity as that value to which one approaches closer and closer as increasing numbers of measurements are made with increasing care. As such, the true value is an idealization similar to the mathematician’s point with no size or line with no width; like the point or line, it is a useful idealization. I will often denote the true values of measured quantities x, y, \dots , by their corresponding capital letters X, Y, \dots . If the measurements of x are subject to many small random errors but negligible systematic errors, their distribution will be a symmetric bell-shaped curve centered on the true value X .

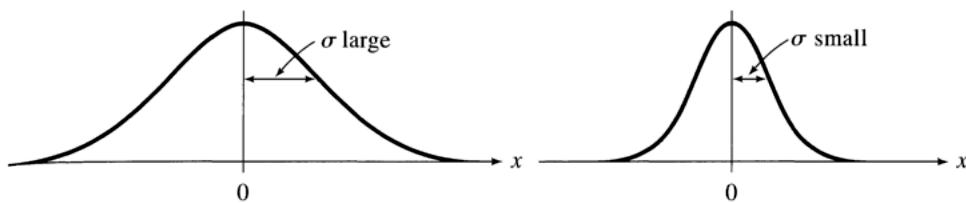


Figure 5.8. The Gauss function (5.17) is bell-shaped and centered on $x = 0$. The bell curve is wide if σ is large and narrow if σ is small. Although for now we will view σ as just a parameter that characterizes the bell curve's width, σ can be shown (as in Problem 5.13) to be the distance from the center of the curve to the point where the curvature changes sign. This distance is shown in the two graphs.

The mathematical function that describes the bell-shaped curve is called the *normal distribution*, or *Gauss function*.² The prototype of this function is

$$e^{-x^2/2\sigma^2}, \quad (5.17)$$

where σ is a fixed parameter I will call the *width parameter*. It is important that you become familiar with the properties of this function.

When $x = 0$, the Gauss function (5.17) is equal to one. The function is symmetric about $x = 0$, because it has the same value for x and $-x$. As x moves away from zero in either direction, $x^2/2\sigma^2$ increases, quickly if σ is small, more slowly if σ is large. Therefore, as x moves away from the origin, the function (5.17) decreases toward zero. Thus, the general appearance of the Gauss function (5.17) is as shown in Figure 5.8. The graphs illustrate the name “width parameter” for σ because the bell shape is wide if σ is large and narrow if σ is small.

The Gauss function (5.17) is a bell-shaped curve centered on $x = 0$. To obtain a bell-shaped curve centered on some other point $x = X$, we merely replace x in (5.17) by $x - X$. Thus, the function

$$e^{-(x-X)^2/2\sigma^2} \quad (5.18)$$

has its maximum at $x = X$ and falls off symmetrically on either side of $x = X$, as in Figure 5.9.

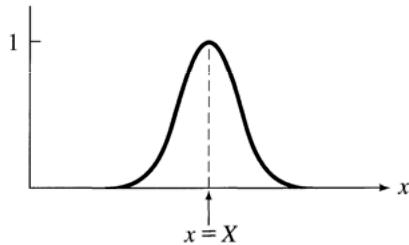


Figure 5.9. The Gauss function (5.18) is bell-shaped and centered on $x = X$.

²Other common names for the Gauss function are the Gaussian function (or just “Gaussian”), the normal density function, and the normal error function. The last of these names is rather unfortunate because the name “error function” is often used for the integral of the Gauss function (as discussed in Section 5.4).

The function (5.18) is not quite in its final form to describe a limiting distribution because any distribution must be *normalized*; that is, it must satisfy

$$\int_{-\infty}^{\infty} f(x) dx = 1. \quad (5.19)$$

To arrange this normalization, we set

$$f(x) = Ne^{-(x-X)^2/2\sigma^2}. \quad (5.20)$$

(Multiplication by the factor N does not change the shape, nor does it shift the maximum at $x = X$.) We must then choose the “normalization factor” N so that $f(x)$ is normalized as in (5.19). This involves some elementary manipulation of integrals, which I give in some detail:

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^{\infty} Ne^{-(x-X)^2/2\sigma^2} dx. \quad (5.21)$$

In evaluating this kind of integral, changing variables to simplify the integral is always a good idea. Thus, we can set $x - X = y$ (in which case $dx = dy$) and get

$$= N \int_{-\infty}^{\infty} e^{-y^2/2\sigma^2} dy. \quad (5.22)$$

Next, we can set $y/\sigma = z$ (in which case $dy = \sigma dz$) and get

$$= N\sigma \int_{-\infty}^{\infty} e^{-z^2/2} dz. \quad (5.23)$$

The remaining integral is one of the standard integrals of mathematical physics. It can be evaluated by elementary methods, but the details are not especially illuminating, so I will simply quote the result;³

$$\int_{-\infty}^{\infty} e^{-z^2/2} dz = \sqrt{2\pi}. \quad (5.24)$$

Returning to (5.21) and (5.23), we find that

$$\int_{-\infty}^{\infty} f(x) dx = N\sigma\sqrt{2\pi}.$$

Because this integral must equal 1, we must choose the normalization factor N to be

$$N = \frac{1}{\sigma\sqrt{2\pi}}.$$

With this choice for the normalization factor, we arrive at the final form for the Gauss, or normal, distribution function, which we denote by $G_{X,\sigma}(x)$:

³For a derivation, see, for example, H. D. Young, *Statistical Treatment of Experimental Data* (McGraw-Hill, 1962), Appendix D.

The Gauss, or Normal, Distribution

$$G_{X,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-X)^2/2\sigma^2}. \quad (5.25)$$

Notice that I have added subscripts X and σ to indicate the center and width of the distribution. The function $G_{X,\sigma}(x)$ describes the limiting distribution of results in a measurement of a quantity x whose true value is X , if the measurement is subject only to random errors. Measurements whose limiting distribution is given by the Gauss function (5.25) are said to be *normally distributed*.

The significance of the width parameter σ will be explored shortly. We have already seen that a small value of σ gives a sharply peaked distribution, which corresponds to a precise measurement, whereas a large value of σ gives a broad distribution, which corresponds to a low-precision measurement. Figure 5.10 shows two examples of Gauss distributions with different centers X and widths σ . Note

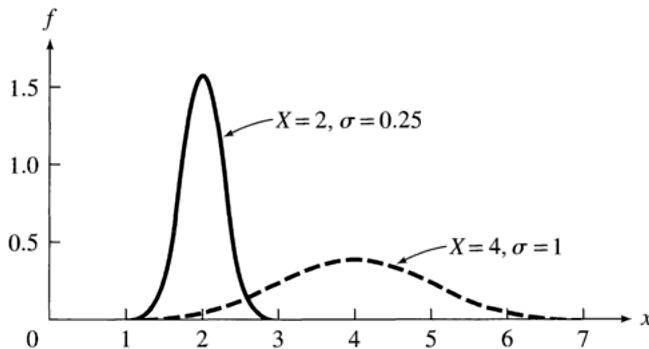


Figure 5.10. Two normal, or Gauss, distributions.

how the factor σ in the denominator of (5.25) guarantees that a narrower distribution (σ smaller) is automatically higher at its center, as it must be so that the total area under the curve equals 1.

Quick Check 5.3. On square-ruled paper, sketch the Gauss function $G_{X,\sigma}(x)$ for $X = 10$ and $\sigma = 1$. Use your calculator to find the values at $x = 10, 10.5, 11, 11.5, 12$, and 12.5 . You don't need to calculate the values for $x < 10$ because you know the function is symmetric about $x = 10$.

We saw in Section 5.2 that knowledge of the limiting distribution for a measurement lets us compute the average value \bar{x} expected after numerous trials. According

to (5.15), this expected average for the Gauss distribution $f(x) = G_{X,\sigma}(x)$ is

$$\bar{x} = \int_{-\infty}^{\infty} x G_{X,\sigma}(x) dx. \quad (5.26)$$

Before we evaluate this integral, we should note that the answer almost obviously will be X , because the symmetry of the Gauss function about X implies that the same number of results will fall any distance above X as will fall an equal distance below X . Thus, the average should be X .

We can calculate the integral (5.26) for the Gauss distribution as follows:

$$\begin{aligned} \bar{x} &= \int_{-\infty}^{\infty} x G_{X,\sigma}(x) dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-(x-X)^2/2\sigma^2} dx. \end{aligned} \quad (5.27)$$

If we make the change of variables $y = x - X$, then $dx = dy$ and $x = y + X$. Thus, the integral (5.27) becomes two terms,

$$\bar{x} = \frac{1}{\sigma\sqrt{2\pi}} \left(\int_{-\infty}^{\infty} y e^{-y^2/2\sigma^2} dy + X \int_{-\infty}^{\infty} e^{-y^2/2\sigma^2} dy \right). \quad (5.28)$$

The first integral here is exactly zero because the contribution from any point y is exactly canceled by that from the point $-y$. The second integral is the normalization integral encountered in (5.22) and has the value $\sigma\sqrt{2\pi}$. This integral cancels with the $\sigma\sqrt{2\pi}$ in the denominator and leaves the expected answer that

$$\bar{x} = X, \quad (5.29)$$

after many trials. In other words, if the measurements are distributed according to the Gauss distribution $G_{X,\sigma}(x)$, then, after numerous trials, the mean value \bar{x} is the true value X , on which the Gauss function is centered.

The result (5.29) would be exactly true only if we could make infinitely many measurements. Its practical usefulness is that if we make a large (but finite) number of trials, our average will be *close* to X .

Another interesting quantity to compute is the *standard deviation* σ_x after a large number of trials. According to (5.16), this quantity is given by

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 G_{X,\sigma}(x) dx. \quad (5.30)$$

This integral is evaluated easily. We replace \bar{x} by X , make the substitutions $x - X = y$ and $y/\sigma = z$, and finally integrate by parts to obtain the result (see Problem 5.16)

$$\sigma_x^2 = \sigma^2 \quad (5.31)$$

after many trials. In other words, the width parameter σ of the Gauss function $G_{X,\sigma}(x)$ is just the standard deviation we would obtain after making many measurements. This is, of course, why the letter σ is used for the width parameter and explains why σ is often called the standard deviation of the Gauss distribution $G_{X,\sigma}(x)$. Strictly speaking, however, σ is the standard deviation expected only after

infinitely many trials. If we make some finite number of measurements (10 or 20, say) of x , the observed standard deviation should be some approximation to σ , but we have no reason to think it will be *exactly* σ . Section 5.5 addresses what more can be said about the mean and standard deviation after a finite number of trials.

5.4 The Standard Deviation as 68% Confidence Limit

The limiting distribution $f(x)$ for measurement of some quantity x tells us the probability of obtaining any given value of x . Specifically, the integral

$$\int_a^b f(x) dx$$

is the probability that any one measurement gives an answer in the range $a \leq x \leq b$. If the limiting distribution is the Gauss function $G_{X,\sigma}(x)$, this integral can be evaluated. In particular, we can now calculate the probability (discussed in Chapter 4) that a measurement will fall within one standard deviation σ of the true value X . This probability is

$$Prob(\text{within } \sigma) = \int_{X-\sigma}^{X+\sigma} G_{X,\sigma}(x) dx \quad (5.32)$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{X-\sigma}^{X+\sigma} e^{-(x-X)^2/2\sigma^2} dx. \quad (5.33)$$

This integral is illustrated in Figure 5.11. It can be simplified in the now familiar

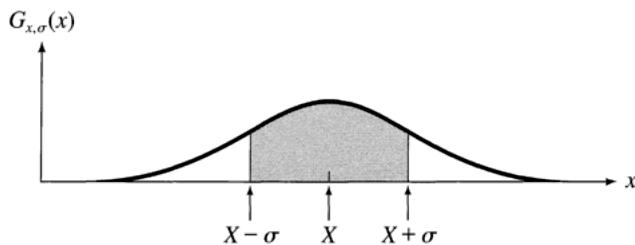


Figure 5.11. The shaded area between $X \pm \sigma$ is the probability of a measurement within one standard deviation of X .

way by substituting $(x - X)/\sigma = z$. With this substitution, $dx = \sigma dz$, and the limits of integration become $z = \pm 1$. Therefore,

$$Prob(\text{within } \sigma) = \frac{1}{\sqrt{2\pi}} \int_{-1}^1 e^{-z^2/2} dz. \quad (5.34)$$

Before discussing the integral (5.34), let me remark that we could equally have found the probability for an answer within 2σ of X or within 1.5σ of X . More generally, we could calculate $Prob(\text{within } t\sigma)$, which means “the probability for an

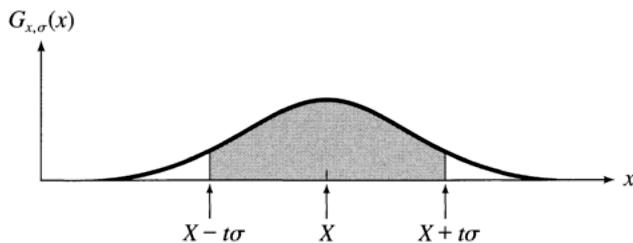


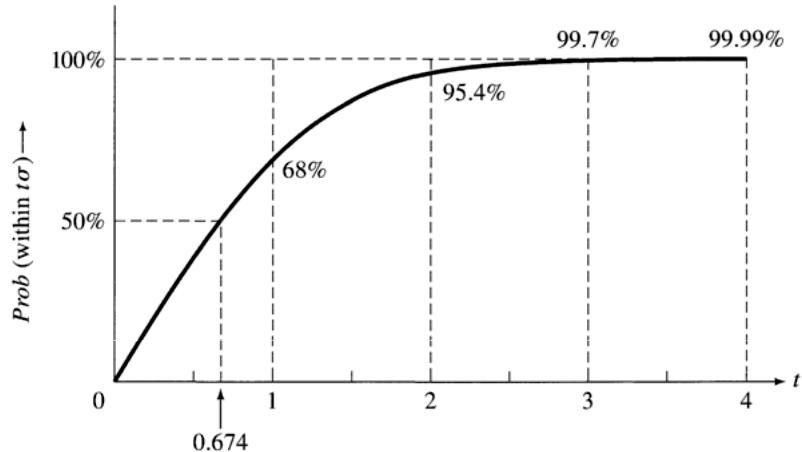
Figure 5.12. The shaded area between $X \pm t\sigma$ is the probability of a measurement within t standard deviations of X .

answer within $t\sigma$ of X ,” where t is any positive number. This probability is given by the area in Figure 5.12, and a calculation identical to that leading to (5.34) gives (as in Problem 5.22)

$$\text{Prob}(\text{within } t\sigma) = \frac{1}{\sqrt{2\pi}} \int_{-t}^t e^{-z^2/2} dz. \quad (5.35)$$

The integral (5.35) is a standard integral of mathematical physics, and it is often called the *error function*, denoted $\text{erf}(t)$, or the *normal error integral*. It cannot be evaluated analytically but is easily calculated on a computer (or even many calculators). Figure 5.13 shows this integral plotted as a function of t and tabulates a few values. A more complete tabulation can be found in Appendix A (see also Appendix B, which shows a different, but closely related, integral).

We first note from Figure 5.13 that the probability that a measurement will fall



t	0	0.25	0.5	0.75	1.0	1.25	1.5	1.75	2.0	2.5	3.0	3.5	4.0
$\text{Prob } (\%)$	0	20	38	55	68	79	87	92	95.4	98.8	99.7	99.95	99.99

Figure 5.13. The probability $\text{Prob}(\text{within } t\sigma)$ that a measurement of x will fall within t standard deviations of the true value $x = X$. Two common names for this function are the *normal error integral* and the *error function*, $\text{erf}(t)$.

within one standard deviation of the true answer is 68%, as anticipated in Chapter 4. If we quote the standard deviation as our uncertainty in such a measurement (that is, write $x = x_{\text{best}} \pm \delta x$, and take $\delta x = \sigma$), then we can be 68% confident that we are within σ of the correct answer.

We can also see in Figure 5.13 that the probability $\text{Prob}(\text{within } t\sigma)$ rapidly approaches 100% as t increases. The probability that a measurement will fall inside 2σ is 95.4%; that for 3σ is 99.7%. To put these results another way, the probability that a measurement will fall *outside* one standard deviation is appreciable (32%), that it will lie outside 2σ is much smaller (4.6%), and that it will lie outside 3σ is extremely small (0.3%).

Of course, nothing is sacred about the number 68%; it just happens to be the confidence associated with the standard deviation σ . One alternative to the standard deviation is called the *probable error*, or PE, and is defined as that distance for which there is a 50% probability of a measurement between $X \pm \text{PE}$. Figure 5.13 shows that (for a measurement that is normally distributed) the probable error is

$$\text{PE} \approx 0.67\sigma.$$

Some experimenters like to quote the PE as the uncertainty in their measurements. Nonetheless, the standard deviation σ is the most popular choice because its properties are so simple.

Quick Check 5.4. The measurements of a certain distance x are distributed normally with $X = 10$ and $\sigma = 2$. What is the probability that a single measurement will lie between $x = 7$ and $x = 13$? What is the probability that it will lie *outside* the range from $x = 7$ to 13?

5.5 Justification of the Mean as Best Estimate

The past three sections have discussed the *limiting distribution* $f(x)$, the distribution obtained from an infinite number of measurements of a quantity x . If $f(x)$ were known, we could calculate the mean \bar{x} and standard deviation σ obtained after infinitely many measurements, and (at least for the normal distribution) we would also know the true value X . Unfortunately, we never do know the limiting distribution. In practice, we have a finite number of measured values (5, 10, or perhaps 50),

$$x_1, x_2, \dots, x_N,$$

and our problem is to arrive at *best estimates* of X and σ based on these N measured values.

If the measurements follow a normal distribution $G_{X,\sigma}(x)$ and if we knew the parameters X and σ , we could calculate the probability of obtaining the values x_1, \dots, x_N that were actually obtained. Thus, the probability of getting a reading

near x_1 , in a small interval dx_1 , is

$$\text{Prob}(x \text{ between } x_1 \text{ and } x_1 + dx_1) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x_1 - X)^2/2\sigma^2} dx_1.$$

In practice, we are not interested in the size of the interval dx_1 (or the factor $\sqrt{2\pi}$), so we abbreviate this equation to

$$\text{Prob}(x_1) \propto \frac{1}{\sigma} e^{-(x_1 - X)^2/2\sigma^2}. \quad (5.36)$$

I will refer to (5.36) as the probability of getting the value x_1 , although strictly speaking it is the probability of getting a value in an interval near x_1 as in the preceding equation.

The probability of obtaining the second reading x_2 is

$$\text{Prob}(x_2) \propto \frac{1}{\sigma} e^{-(x_2 - X)^2/2\sigma^2}, \quad (5.37)$$

and we can similarly write down all the probabilities ending with

$$\text{Prob}(x_N) \propto \frac{1}{\sigma} e^{-(x_N - X)^2/2\sigma^2}. \quad (5.38)$$

Equations (5.36) through (5.38) give the probabilities of obtaining each of the readings x_1, \dots, x_N , calculated in terms of the assumed limiting distribution $G_{X,\sigma}(x)$. The probability that we observe the whole set of N readings is just the product of these separate probabilities,⁴

$$\text{Prob}_{X,\sigma}(x_1, \dots, x_N) = \text{Prob}(x_1) \times \text{Prob}(x_2) \times \dots \times \text{Prob}(x_N)$$

or

$$\text{Prob}_{X,\sigma}(x_1, \dots, x_N) \propto \frac{1}{\sigma^N} e^{-\sum(x_i - X)^2/2\sigma^2}. \quad (5.39)$$

Understanding the significance of the various quantities in (5.39) is most important. The numbers x_1, \dots, x_N are the actual results of N measurements; thus x_1, \dots, x_N are known, fixed numbers. The quantity $\text{Prob}_{X,\sigma}(x_1, \dots, x_N)$ is the probability for obtaining the N results x_1, \dots, x_N , calculated in terms of X and σ , the true value of x and the width parameter of its distribution. The numbers X and σ are *not* known; we want to find best estimates for X and σ based on the given observations x_1, \dots, x_N . I have added subscripts X and σ to the probability (5.39) to emphasize that it depends on the (unknown) values of X and σ .

Because the actual values of X and σ are unknown, we might imagine guessing values X' and σ' and then using those guessed values to compute the probability $\text{Prob}_{X',\sigma'}(x_1, \dots, x_N)$. If we next guessed two new values, X'' and σ'' and found that the corresponding probability $\text{Prob}_{X'',\sigma''}(x_1, \dots, x_N)$ was larger, we would natu-

⁴We are using the well-known result that the probability for several independent events is the product of their separate probabilities. For example, the probability of throwing a “heads” with a coin is 1/2 and that of throwing a “six” with a die is 1/6. Therefore, the probability of throwing a “heads” *and* a “six” is $(1/2) \times (1/6) = 1/12$.

rally regard the new values X'' and σ'' as better estimates for X and σ . Continuing in this way, we could imagine hunting for the values of X and σ that make $Prob_{X,\sigma}(x_1, \dots, x_N)$ as large as possible, and these values would be regarded as the best estimates for X and σ .

This plausible procedure for finding the best estimates for X and σ is called by statisticians the *principle of maximum likelihood*. It can be stated briefly as follows:

Given the N observed measurements x_1, \dots, x_N , the best estimates for X and σ are those values for which the observed x_1, \dots, x_N are most likely. That is, the best estimates for X and σ are those values for which $Prob_{X,\sigma}(x_1, \dots, x_N)$ is maximum, given that here

$$Prob_{X,\sigma}(x_1, \dots, x_N) \propto \frac{1}{\sigma^N} e^{-\sum(x_i-X)^2/2\sigma^2}. \quad (5.40)$$

Using this principle, we can easily find the best estimate for the true value X . Obviously (5.40) is *maximum* if the sum in the exponent is *minimum*. Thus, the best estimate for X is that value of X for which

$$\sum_{i=1}^N (x_i - X)^2 / \sigma^2 \quad (5.41)$$

is minimum. To locate this minimum, we differentiate with respect to X and set the derivative equal to zero, giving

$$\sum_{i=1}^N (x_i - X) = 0$$

or

$$\text{(best estimate for } X) = \frac{\sum x_i}{N}. \quad (5.42)$$

That is, the *best estimate* for the true value X is the mean of our N measurements, $\bar{x} = \sum x_i / N$, a result we have been assuming without proof since Chapter 1.

Finding the best estimate for σ , the width of the limiting distribution, is a little harder, because the probability (5.40) is a more complicated function of σ . We must differentiate (5.40) with respect to σ and set the derivative equal to zero. (I leave the details to you; see Problem 5.26.) This procedure gives the value of σ that maximizes (5.40) and that is therefore the best estimate for σ , as

$$\text{(best estimate for } \sigma) = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - X)^2}. \quad (5.43)$$

The true value X is unknown. Thus, in practice, we have to replace X in (5.43) by our best estimate for X , namely the mean \bar{x} . This replacement yields the estimate

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}. \quad (5.44)$$

In other words, our estimate for the width σ of the limiting distribution is the standard deviation of the N observed values, x_1, \dots, x_N , as originally defined in (4.6).

You may have been surprised that the estimate (5.44) is the same as our original definition (4.6), using N , of the standard deviation, instead of our “improved” definition, using $N - 1$. In fact, in passing from the best estimate (5.43) to the expression (5.44), we have glossed over a rather elegant subtlety. The best estimate (5.43) involves the true value X , whereas in (5.44) we have replaced X by \bar{x} (our best estimate for X). Now, these numbers are generally not the same, and you can easily see that the number (5.44) is *always less than*, or at most equal to, (5.43).⁵ Thus, in passing from (5.43) to (5.44), we have consistently *underestimated* the width σ . Appendix E shows that this underestimation is corrected by replacing the denominator N in (5.44) by $N - 1$. That is, the best estimate for the width σ is precisely the “improved,” or “sample,” standard deviation of the measured values x_1, x_2, \dots, x_N , with $(N - 1)$ in its denominator,

$$\text{(best estimate for } \sigma) = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}. \quad (5.45)$$

Now is a good time to review the rather complicated story that has unfolded so far. First, if the measurements of x are subject only to random errors, their limiting distribution is the Gauss function $G_{X,\sigma}(x)$ centered on the true value X and with width σ . The width σ is the 68% confidence limit, in that there is a 68% probability that any measurement will fall within a distance σ of the true value X . In practice, neither X nor σ is known. Instead, we know our N measured values x_1, \dots, x_N , where N is as large as our time and patience allowed us to make it. Based on these N measured values, our best estimate of the true value X has been shown to be the mean $\bar{x} = \sum x_i / N$, and our best estimate of the width σ is the standard deviation σ_x of x_1, \dots, x_N as defined in (5.45).

Two further questions now arise. First, what is the uncertainty in \bar{x} as an estimate of the true value of X ? This question is discussed in Section 5.7, where the uncertainty in \bar{x} is shown to be the *standard deviation of the mean*, or SDOM, as defined in Chapter 4. Second, what is the uncertainty in σ_x as an estimate of the true width σ ? The formula for this “uncertainty in the uncertainty” or “standard deviation of the standard deviation” is derived in Appendix E; the result proved there is that the fractional uncertainty in σ_x is

$$\text{(fractional uncertainty in } \sigma_x) = \frac{1}{\sqrt{2(N-1)}}. \quad (5.46)$$

⁵If we regard (5.43) as a function of X , we have just seen that this function is minimum at $X = \bar{x}$. Thus (5.44) is always less than or equal to (5.43).

This result makes clear the need for numerous measurements before the uncertainty can be known reliably. For example, with just three measurements of a quantity ($N = 3$), the result (5.46) implies that the standard deviation is 50% uncertain!

Quick Check 5.5. To test the reliability of a ballistic galvanometer, a student discharges a capacitor (charged to a fixed voltage) through the galvanometer three times and measures the resulting charge q in microcoulombs. From his three measurements, he calculates the standard deviation to be $\sigma_q = 6 \mu\text{C}$. Use (5.46) to find the uncertainty $\delta\sigma_q$ in this parameter. The true value of the standard deviation could easily be as small as $\sigma_q - \delta\sigma_q$ or as large as $\sigma_q + \delta\sigma_q$. What are these two values for this experiment? Note well how unreliable σ_q is after only three measurements.

The results of the past two sections depend on the assumption that our measurements are normally distributed.⁶ Although this assumption is reasonable, it is difficult to verify in practice and is sometimes not exactly true. This being the case, I should emphasize that, even when the distribution of measurements is *not* normal, it is almost always *approximately* normal, and you can safely use the ideas of this chapter, at least as good approximations.

5.6 Justification of Addition in Quadrature

Let us now return to the topic of Chapter 3, the propagation of errors. I stated there, without formal proof, that when errors are random and independent, they can be combined in quadrature according to certain standard rules, either the “simple rules” in (3.16) and (3.18), or the general rule in (3.47), which includes the “simple” rules as special cases. This use of addition in quadrature can now be justified.

The problem of error propagation arises when we measure one or more quantities x, \dots, z , all with uncertainties, and then use our measured values to calculate some quantity $q(x, \dots, z)$. The main question is, of course, to decide on the uncertainty in our answer for q . If the quantities x, \dots, z are subject only to random errors, they will be normally distributed with width parameters⁷ $\sigma_x, \dots, \sigma_z$, which we take to be the uncertainties associated with any single measurement of the corresponding quantities. The question to be decided now is this: Knowing the distributions of measurements of x, \dots, z , what can we say about the distribution of values for q ? In particular, what will be the width of the distribution of values of q ? This question is answered in four steps, numbered I to IV.

⁶And that systematic errors have been reduced to a negligible level.

⁷When discussing several different measured quantities x, \dots, z , I use subscripts x, \dots, z to distinguish the width parameters of their limiting distributions. Thus σ_x denotes the width of the Gauss distribution $G_{X,\sigma_x}(x)$ for the measurements of x , and so on.

I. MEASURED QUANTITY PLUS FIXED NUMBER

First, consider two simple, special cases. Suppose we measure a quantity x and proceed to calculate the quantity

$$q = x + A, \quad (5.47)$$

where A is a fixed number with no uncertainty (such as $A = 1$ or π). Suppose also that the measurements of x are normally distributed about the true value X , with

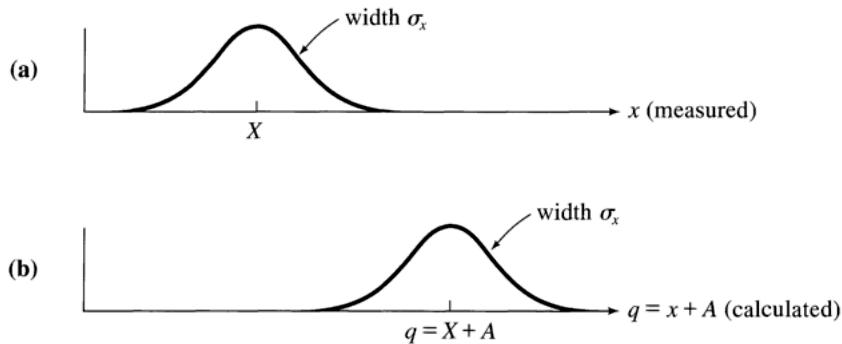


Figure 5.14. If the measured values of x are normally distributed with center $x = X$ and width σ_x , the calculated values of $q = x + A$ (with A fixed and known) will be normally distributed with center $q = X + A$ and the same width σ_x .

width σ_x , as in Figure 5.14(a). Then the probability of obtaining any value x (in a small interval dx) is $G_{X,\sigma_x}(x) dx$ or

$$(\text{probability of obtaining value } x) \propto e^{-(x-X)^2/2\sigma_x^2}. \quad (5.48)$$

Our problem is to deduce the probability of obtaining any value q of the quantity defined by (5.47). Now, from (5.47) we see that $x = q - A$ and hence that

$$(\text{probability of obtaining value } q) = (\text{probability of obtaining } x = q - A).$$

The second probability is given by (5.48), and so

$$\begin{aligned} (\text{probability of obtaining value } q) &\propto e^{-[(q-A)-X]^2/2\sigma_x^2} \\ &= e^{-[q-(X+A)]^2/2\sigma_x^2}. \end{aligned} \quad (5.49)$$

The result (5.49) shows that the calculated values of q are normally distributed and centered on the value $X + A$, with width σ_x , as shown in Figure 5.14(b). In particular, the uncertainty in q is the same (namely, σ_x) as that in x , just as the rule (3.16) would have predicted.

II. MEASURED QUANTITY TIMES FIXED NUMBER

As a second simple example, suppose that we measure x and calculate the quantity

$$q = Bx,$$

where B is a fixed number (such as $B = 2$ or $B = \pi$). If the measurements of x are normally distributed, then, arguing exactly as before, we conclude that⁸

$$\begin{aligned} (\text{probability of obtaining value } q) &\propto (\text{probability of obtaining } x = q/B) \\ &\propto \exp\left[-\left(\frac{q}{B} - X\right)^2/2\sigma_x^2\right] \\ &= \exp[-(q - BX)^2/2B^2\sigma_x^2]. \end{aligned} \quad (5.50)$$

In other words, the values of $q = Bx$ will be normally distributed, with center at $q = BX$ and width $B\sigma_x$, as shown in Figure 5.15. In particular, the uncertainty in $q = Bx$ is B times that in x , just as our rule (3.18) would have predicted.

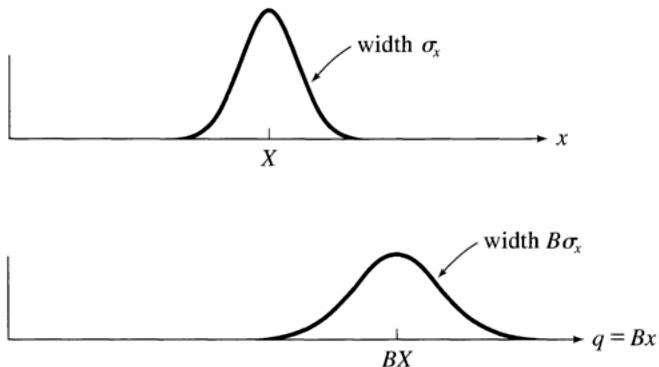


Figure 5.15. If the measured values of x are normally distributed with center $x = X$ and width σ_x , then the calculated values of $q = Bx$ (with B fixed and known) will be normally distributed with center BX and width $B\sigma_x$.

III. SUM OF TWO MEASURED QUANTITIES

As a first nontrivial example of error propagation, suppose we measure two independent quantities x and y and calculate their sum $x + y$. We suppose that the measurements of x and y are normally distributed about their true values X and Y , with widths σ_x and σ_y as in Figures 5.16(a) and (b), and we will try to find the distribution of the calculated values of $x + y$. We will find that the values of $x + y$

⁸Here I introduce the alternative notation $\exp(z)$ for the exponential function, $\exp(z) \equiv e^z$. When the exponent z becomes complicated, the “exp” notation is more convenient to type or print.

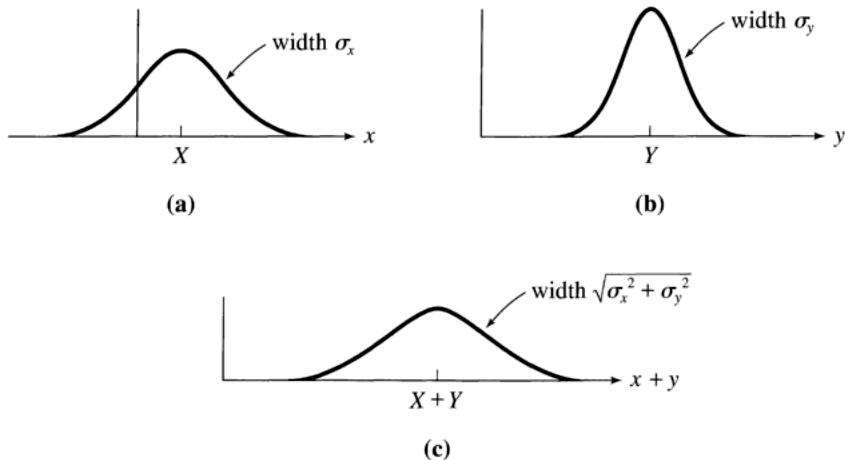


Figure 5.16. If the measurements of x and y are independent and normally distributed with centers X and Y and widths σ_x and σ_y , then the calculated values of $x+y$ are normally distributed with center $X+Y$ and width $\sqrt{\sigma_x^2 + \sigma_y^2}$.

are normally distributed, that their center is the true value $X + Y$, and that the width of their distribution is

$$\sqrt{\sigma_x^2 + \sigma_y^2},$$

as in Figure 5.16(c). In particular, this result justifies the rule of Chapter 3 that if x and y are subject to independent random uncertainties only, then the uncertainty in $x + y$ is the quadratic sum of the separate uncertainties in x and y .

To simplify our algebra, we assume at first that the true values X and Y are both zero. In this case, the probability of getting any particular value of x is

$$Prob(x) \propto \exp\left(\frac{-x^2}{2\sigma_x^2}\right) \quad (5.51)$$

and that of y is

$$Prob(y) \propto \exp\left(\frac{-y^2}{2\sigma_y^2}\right). \quad (5.52)$$

Our problem now is to calculate the probability of obtaining any particular value of $x + y$. We first observe that because x and y are independently measured, the probability of obtaining any given x and any given y is just the product of (5.51) and (5.52):

$$Prob(x, y) \propto \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right]. \quad (5.53)$$

Knowing the probability of obtaining any x and any y , we can now calculate the probability for any given value of $x + y$. The first step is to rewrite the exponent in (5.53) in terms of the variable of interest, $x + y$. This step can be done using the

identity (which you can easily verify)

$$\frac{x^2}{A} + \frac{y^2}{B} = \frac{(x+y)^2}{A+B} + \frac{(Bx-Ay)^2}{AB(A+B)} \quad (5.54)$$

$$= \frac{(x+y)^2}{A+B} + z^2. \quad (5.55)$$

In the second line I have introduced the abbreviation z^2 for the second term on the right of (5.54) because its value does not interest us anyway.

If we substitute (5.55) into (5.53), replacing A with σ_x^2 and B with σ_y^2 , we obtain

$$Prob(x, y) \propto \exp\left[-\frac{(x+y)^2}{2(\sigma_x^2 + \sigma_y^2)} - \frac{z^2}{2}\right]. \quad (5.56)$$

This probability for obtaining given values of x and y can just as well be viewed as the probability of obtaining given values of $x+y$ and z . Thus, we can rewrite (5.56) as

$$Prob(x+y, z) \propto \exp\left[\frac{-(x+y)^2}{2(\sigma_x^2 + \sigma_y^2)}\right] \exp\left[\frac{-z^2}{2}\right]. \quad (5.57)$$

Finally, what we want is the probability of obtaining a given value of $x+y$, *irrespective of the value of z* . This probability is obtained by summing, or rather integrating, (5.57) over all possible values of z ; that is,

$$Prob(x+y) = \int_{-\infty}^{\infty} Prob(x+y, z) dz. \quad (5.58)$$

When we integrate (5.57) with respect to z , the factor $\exp(-z^2/2)$ integrates to $\sqrt{2\pi}$, and we find

$$Prob(x+y) \propto \exp\left[\frac{-(x+y)^2}{2(\sigma_x^2 + \sigma_y^2)}\right]. \quad (5.59)$$

This result shows that the values of $x+y$ are normally distributed with width $\sqrt{\sigma_x^2 + \sigma_y^2}$ as anticipated.

Our proof is complete for the case when the true values of x and y are both zero, $X = Y = 0$. If X and Y are nonzero, we can proceed as follows: We first write

$$x+y = (x-X) + (y-Y) + (X+Y). \quad (5.60)$$

Here, the first two terms are centered on zero, with widths σ_x and σ_y , by the result from step I. Therefore, by the result just proved [Equation (5.59)], the sum of the first two terms is normally distributed with width $\sqrt{\sigma_x^2 + \sigma_y^2}$. The third term in (5.60) is a fixed number; therefore, by the result of step I again, it simply shifts the center of the distribution to $(X+Y)$ but leaves the width unchanged. In other words, the values of $(x+y)$ as given by (5.60) are normally distributed about $(X+Y)$ with width $\sqrt{\sigma_x^2 + \sigma_y^2}$. This is the required result.

IV. THE GENERAL CASE

Having justified the error-propagation formula for the special case of a sum $x + y$, we can justify the general formula for error propagation surprisingly simply. Suppose we measure two independent quantities x and y whose observed values are normally distributed, and we now calculate some quantity $q(x, y)$ in terms of x and y . The distribution of values of $q(x, y)$ is found easily by using the results from steps I through III as follows:

First, the widths σ_x and σ_y (the uncertainties in x and y) must, as always, be small. This requirement means that we are concerned only with values of x close to X and y close to Y , and we can use the approximation (3.42) to write

$$q(x, y) \approx q(X, Y) + \left(\frac{\partial q}{\partial x} \right) (x - X) + \left(\frac{\partial q}{\partial y} \right) (y - Y). \quad (5.61)$$

This approximation is good because the only values of x and y that occur significantly often are close to X and Y . The two partial derivatives are evaluated at X and Y and are, therefore, fixed numbers.

The approximation (5.61) expresses the desired quantity $q(x, y)$ as the sum of three terms. The first term $q(X, Y)$ is a fixed number, so it merely shifts the distribution of answers. The second term is the fixed number $\partial q / \partial x$ times $(x - X)$, whose distribution has width σ_x , so the values of the second term are centered on zero, with width

$$\left(\frac{\partial q}{\partial x} \right) \sigma_x.$$

Similarly, the values of the third term are centered on zero with width

$$\left(\frac{\partial q}{\partial y} \right) \sigma_y.$$

Combining the three terms in (5.61) and invoking the results already established, we conclude that the values of $q(x, y)$ are normally distributed about the true value $q(X, Y)$ with width

$$\sigma_q = \sqrt{\left(\frac{\partial q}{\partial x} \sigma_x \right)^2 + \left(\frac{\partial q}{\partial y} \sigma_y \right)^2}. \quad (5.62)$$

If we identify the standard deviations σ_x and σ_y as the uncertainties in x and y , the result (5.62) is precisely the rule (3.47) for propagation of random errors, for the case when q is a function of just two variables, $q(x, y)$. If q depends on several variables, $q(x, y, \dots, z)$, the preceding argument can be extended immediately to establish the general rule (3.47) for functions of several variables. Because the rules of Chapter 3 concerning propagation of random errors can be derived from (3.47), they are all now justified.

5.7 Standard Deviation of the Mean

One more important result, quoted in Chapter 4, remains to be proved. This result concerns the standard deviation of the mean $\sigma_{\bar{x}}$. I proved (in Section 5.5) that if we make N measurements x_1, \dots, x_N of a quantity x (that is normally distributed), the best estimate of the true value X is the mean \bar{x} of x_1, \dots, x_N . In Chapter 4, I stated that the uncertainty in this estimate is the standard deviation of the mean,

$$\sigma_{\bar{x}} = \sigma_x / \sqrt{N}. \quad (5.63)$$

Let us now prove this result. The proof is so surprisingly brief that you need to follow it very carefully.

Suppose that the measurements of x are normally distributed about the true value X with width parameter σ_x . We now want to know the reliability of *the average of the N measurements*. To answer this, we naturally imagine repeating our N measurements many times; that is, we imagine performing a sequence of experiments, in each of which we make N measurements and compute the average. We now want to find the distribution of these many determinations of the average of N measurements.

In each experiment, we measure N quantities x_1, \dots, x_N and then compute the function

$$\bar{x} = \frac{x_1 + \dots + x_N}{N}. \quad (5.64)$$

Because the calculated quantity (\bar{x}) is a simple function of the measured quantities x_1, \dots, x_N , we can now find the distribution of our answers for \bar{x} by using the error-propagation formula. The only unusual feature of the function (5.64) is that all the measurements x_1, \dots, x_N happen to be measurements of the same quantity, with the same true value X and the same width σ_x .

We first observe that, because each of the measured quantities x_1, \dots, x_N is normally distributed, the same is true for the function \bar{x} given by (5.64). Second, the true value for each of x_1, \dots, x_N is X ; so the true value of \bar{x} as given by (5.64) is

$$\frac{X + \dots + X}{N} = X.$$

Thus, after making many determinations of the average \bar{x} of N measurements, our many results for \bar{x} will be normally distributed about the true value X . The only remaining (and most important) question is to find the width of our distribution of answers. According to the error-propagation formula (5.62), rewritten for N variables, this width is

$$\sigma_{\bar{x}} = \sqrt{\left(\frac{\partial \bar{x}}{\partial x_1} \sigma_{x_1}\right)^2 + \dots + \left(\frac{\partial \bar{x}}{\partial x_N} \sigma_{x_N}\right)^2}. \quad (5.65)$$

Because x_1, \dots, x_N are all measurements of the same quantity x , their widths are all the same and are all equal to σ_x ,

$$\sigma_{x_1} = \dots = \sigma_{x_N} = \sigma_x.$$

We also see from (5.64) that all the partial derivatives in (5.65) are the same:

$$\frac{\partial \bar{x}}{\partial x_1} = \dots = \frac{\partial \bar{x}}{\partial x_N} = \frac{1}{N}.$$

Therefore, (5.65) reduces to

$$\begin{aligned}\sigma_{\bar{x}} &= \sqrt{\left(\frac{1}{N} \sigma_x\right)^2 + \dots + \left(\frac{1}{N} \sigma_x\right)^2} \\ &= \sqrt{N \frac{\sigma_x^2}{N^2}} = \frac{\sigma_x}{\sqrt{N}},\end{aligned}\tag{5.66}$$

as required.

We have arrived at the desired result (5.66) so quickly that we probably need to pause and review its significance. We imagined a large number of experiments, in each of which we made N measurements of x and then computed the average \bar{x} of those N measurements. We have shown that, after repeating this experiment many times, our many answers for \bar{x} will be normally distributed, that they will be centered on the true value X , and that the width of their distribution is $\sigma_{\bar{x}} = \sigma_x/\sqrt{N}$, as shown in Figure 5.17 for $N = 10$. This width $\sigma_{\bar{x}}$ is the 68% confidence limit for our experiment. If we find the mean of N measurements *once*, we can be 68% confident that our answer lies within a distance $\sigma_{\bar{x}}$ of the true value X . This result is exactly what should be signified by the *uncertainty in the mean*. It also explains clearly why that uncertainty is called the standard deviation of the mean.

With this simple and elegant proof, all the results quoted in earlier chapters concerning random uncertainties have now been justified.

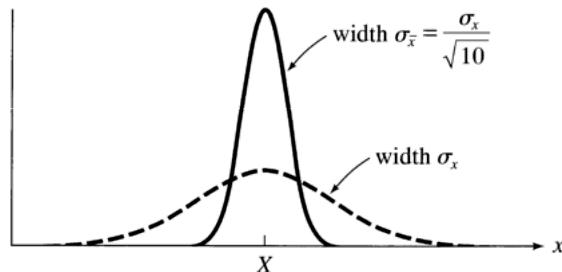


Figure 5.17. The individual measurements of x are normally distributed about X with width σ_x (dashed curve). If we use the same equipment to make many determinations of the average of 10 measurements, the results \bar{x} will be normally distributed about X with width $\sigma_{\bar{x}} = \sigma_x/\sqrt{10}$ (solid curve).

5.8 Acceptability of a Measured Answer

We can now return to two questions first raised but not completely answered in Chapter 2. First, what is meant by the now-familiar statement that we are “reasonably confident” that a measured quantity lies in the range $x_{\text{best}} \pm \delta x$? Second, when we compare our value x_{best} with an expected value x_{exp} (the latter based on some theory, or just someone else’s measurement), how do we decide whether or not the agreement between our value and the expected value is acceptable?

For the first question, the answer should by now be clear: If we measure a quantity x several times (as we usually would), the mean \bar{x} of our measurements is the best estimate for x , and the standard deviation of the mean $\sigma_{\bar{x}}$ is a good measure of its uncertainty. We would report the conclusion that

$$(\text{value of } x) = \bar{x} \pm \sigma_{\bar{x}},$$

meaning that, based on our observations, we expect 68% of any measurements of x , made in the same way, to fall in the range $\bar{x} \pm \sigma_{\bar{x}}$.

We *could* choose to characterize our uncertainty differently. For example, we might choose to state our conclusion as

$$(\text{value of } x) = \bar{x} \pm 2\sigma_{\bar{x}};$$

here, we would be stating the range in which we expect 95% of all comparable measurements to fall. Clearly, the essential point in stating any measured value is to state a range (or uncertainty) and *the confidence level corresponding to that range*. The most popular choice is to give the standard deviation of the answer, with its familiar significance as the 68% confidence limit.

As emphasized in Chapter 2, almost all experimental conclusions involve the comparison of two or more numbers. With our statistical theory, we can now give a quantitative significance to many such comparisons. Here, I will consider just one type of experiment, in which we arrive at a number and compare our result with some known, expected answer. Notice that this general description fits many interesting experiments. For instance, in an experiment to check conservation of momentum, we may measure initial and final momenta, p and p' , to verify that $p = p'$ (within uncertainties), but we can equally well regard this experiment as finding a value for $(p - p')$ to be compared with the expected answer of zero. More generally, when we want to compare any two measurements that are supposedly the same, we can form their difference and compare it with the expected answer of zero. Any experiment that involves measuring a quantity (such as g , the acceleration of gravity), for which an accurate accepted value is known, is also of this type, and the expected answer is the known accepted value.

Let us suppose that a student measures some quantity x (such as the difference of two momenta that are supposedly equal) in the form

$$(\text{value of } x) = x_{\text{best}} \pm \sigma,$$

where σ denotes the standard deviation of his answer (which would be the SDOM if x_{best} was the mean of several measurements). He now wants to compare this answer with the expected answer x_{exp} .

In Chapter 2, I argued that if the discrepancy $|x_{\text{best}} - x_{\text{exp}}|$ is less than (or only slightly more than) σ , then the agreement is satisfactory, but if $|x_{\text{best}} - x_{\text{exp}}|$ is much greater than σ , it is not satisfactory. These criteria are correct as far as they go, but they give no quantitative measure of how good or bad the agreement is. They also do not tell us where to draw the boundary of acceptability. Would a discrepancy of 1.5σ have been satisfactory? Or 2σ ?

We can now answer these questions if we assume that our student's measurement was governed by a normal distribution (as is certainly reasonable). We start by making two working hypotheses about this distribution:

- (1) The distribution is centered on the expected answer x_{exp} .
- (2) The width parameter of the distribution is equal to the student's estimate σ .

Hypothesis (1) is, of course, what the student hopes is true. It assumes that all systematic errors were reduced to a negligible level (so that the distribution was centered on the true value) and that the true value was indeed x_{exp} (that is, that the reasons for expecting x_{exp} were correct). Hypothesis (2) is an approximation because σ must have been an estimate of the standard deviation, but it is a reasonable approximation if the number of measurements on which σ is based is large.⁹ Taken together, our two hypotheses are the same as assuming that the student's procedures and calculations were essentially correct.

We must now decide whether the student's value x_{best} was a reasonable one to obtain if our hypotheses were correct. If the answer is yes, there is no reason to doubt the hypotheses, and all is well; if the answer is no, the hypotheses must be doubted, and the student must examine the possibilities of mistakes in the measurements or calculations, of undetected systematic errors, and of the expected answer x_{exp} being incorrect.

We first determine the discrepancy, $|x_{\text{best}} - x_{\text{exp}}|$, and then

$$t = \frac{|x_{\text{best}} - x_{\text{exp}}|}{\sigma}, \quad (5.67)$$

the number of standard deviations by which x_{best} differs from x_{exp} . (Here, σ denotes the standard deviation appropriate to x_{best} ; if x_{best} is the mean of several measurements, then σ is the standard deviation of the mean.) Next, from the table of the normal error integral in Appendix A, we can find the probability (given our hypotheses) of obtaining an answer that differs from x_{exp} by t or more standard deviations. This probability is

$$\text{Prob}(\text{outside } t\sigma) = 1 - \text{Prob}(\text{within } t\sigma). \quad (5.68)$$

If this probability is large, the discrepancy $|x_{\text{best}} - x_{\text{exp}}|$ is perfectly reasonable, and the result x_{best} is acceptable; if the probability in (5.68) is "unreasonably small," the

⁹We are going to judge the reasonableness of our measurement, x_{best} , by comparing $|x_{\text{best}} - x_{\text{exp}}|$ with σ , our *estimate* of the width of the normal distribution concerned. If the number of measurements on which σ was based is *small*, this estimate may be fairly unreliable, and the confidence levels will be correspondingly inaccurate (although still a useful rough guide). With a small number of measurements, the accurate calculation of confidence limits requires use of the so-called "student's *t* distribution," which allows for the probable variations in our estimate σ of the width. See H. L. Alder and E. B. Roessler, *Introduction to Probability and Statistics*, 6th ed. (W. H. Freeman, 1977), Chapter 10.

discrepancy must be judged *significant* (that is, unacceptable), and our unlucky student must try to find out what has gone wrong.

Suppose, for example, the discrepancy $|x_{\text{best}} - x_{\text{exp}}|$ is one standard deviation. The probability of a discrepancy this large or larger is the familiar 32%. Clearly, a discrepancy of one standard deviation is quite likely to occur and is, therefore, insignificant. At the opposite extreme, the probability $\text{Prob}(\text{outside } 3\sigma)$ is just 0.3%, and, if our hypotheses are correct, a discrepancy of 3σ is most unlikely. Turning this statement around, if our student's discrepancy is 3σ , our hypotheses were most unlikely to be correct.

The boundary between acceptability and unacceptability depends on the level below which we judge a discrepancy is unreasonably improbable. This level is a matter of opinion, to be decided by the experimenter. Many scientists regard 5% a fair boundary for "unreasonable improbability." If we accept this choice, then a discrepancy of 2σ would be just unacceptable, because $\text{Prob}(\text{outside } 2\sigma) = 4.6\%$. In fact, from the table in Appendix A, we see that any discrepancy greater than 1.96σ is unacceptable at this 5% level, and discrepancies this large are sometimes called "*significant*." Similarly, at the 1% level, any discrepancy greater than 2.58σ would be unacceptable, and discrepancies this large are sometimes called "*highly significant*."

Quick Check 5.6. A student measures the electron charge e and notes that her answer is 2.4 standard deviations away from the accepted value. Is this discrepancy significant at the 5% level? What about the 2% level? Or 1%?

We still do not have a clear-cut answer that a certain measured value x_{best} is, or is not, acceptable. Our theory of the normal distribution, however, has given us a clear, quantitative measure of the reasonableness of any particular answer, which is the best we can hope for.

Most physicists do not spend a lot of time debating precisely where the boundary of acceptability lies. If the discrepancy is appreciably less than 2σ (1.8σ , say), then by almost any standard the result would be judged acceptable. If the discrepancy is appreciably more than 2.5σ , then by any standard it is unacceptable. If the discrepancy falls in the gray region between about 1.9σ and about 2.6σ , the experiment is simply inconclusive. If the experiment is sufficiently important (as a test of a new theory, for example), it needs to be repeated (preferably with improved techniques) until a conclusive result *is* obtained.

Of course, some experiments are more complicated and require correspondingly more involved analyses. Most of the basic principles, however, have been illustrated by the simple, important case discussed here. Further examples can be found in Part II.

Principal Definitions and Equations of Chapter 5

LIMITING DISTRIBUTIONS

If $f(x)$ is the limiting distribution for measurement of a continuous variable x , then

$f(x) dx$ = probability that any one measurement will give an answer between x and $x + dx$,

and

$\int_a^b f(x) dx$ = probability that any one measurement will give an answer between $x = a$ and $x = b$. [See (5.12)]

The normalization condition is

$$\int_{-\infty}^{\infty} f(x) dx = 1. \quad [\text{See (5.13)}]$$

The mean value of x expected after many measurements is

$$\bar{x} = \int_{-\infty}^{\infty} x f(x) dx. \quad [\text{See (5.15)}]$$

THE GAUSS, OR NORMAL, DISTRIBUTION

If the measurements of x are subject to many small random errors but negligible systematic error, their limiting distribution will be the *normal*, or *Gauss, distribution*:

$$G_{X,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-X)^2/2\sigma^2}, \quad [\text{See (5.25)}]$$

where

- X = true value of x
- = center of distribution
- = mean value after many measurements,

and

- σ = width parameter of distribution
- = standard deviation after many measurements.

The probability of a single measurement falling within t standard deviations of X is

$$\text{Prob}(\text{within } t\sigma) = \frac{1}{\sqrt{2\pi}} \int_{-t}^t e^{-z^2/2} dz. \quad [\text{See (5.35)}]$$

This integral is often called the error function or the normal error integral. Its value as a function of t is tabulated in Appendix A. In particular,

$$\text{Prob}(\text{within } \sigma) = 68.27\%.$$

ESTIMATING X AND σ FROM N MEASURED VALUES

After N measurements of a normally distributed quantity x ,

$$x_1, x_2, \dots, x_N,$$

the best estimate for the true value X is the mean of our measurements,

$$(\text{best estimate for } X) = \bar{x} = \frac{\sum x_i}{N}, \quad [\text{See (5.42)}]$$

and the best estimate for the width σ is the standard deviation of the measurements,

$$(\text{best estimate for } \sigma) = \sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{(N - 1)}}. \quad [\text{See (5.45)}]$$

The uncertainties in these estimates are as follows: The uncertainty in \bar{x} as an estimate of X is

$$(\text{uncertainty in } \bar{x}) = \text{SDOM} = \frac{\sigma_x}{\sqrt{N}} \quad [\text{See (5.66)}]$$

and the uncertainty in σ_x as the estimate of the true width σ is given by

$$(\text{fractional uncertainty in } \sigma_x) = \frac{1}{\sqrt{2(N - 1)}}. \quad [\text{See (5.46)}]$$

ACCEPTABILITY OF A MEASURED ANSWER

Suppose we measure a quantity x in the standard form

$$(\text{value of } x) = x_{\text{best}} \pm \sigma,$$

where σ is the appropriate standard deviation. Suppose also that, based on some theory or on someone else's measurements, we expected the value x_{exp} . We say that x_{best} differs from x_{exp} by t standard deviations, where

$$t = \frac{|x_{\text{best}} - x_{\text{exp}}|}{\sigma}.$$

Assuming x is normally distributed about x_{exp} with width σ , we can find from Appendix A the probability $\text{Prob}(\text{outside } t\sigma)$ of a discrepancy as large as ours or larger. If this probability is less than some chosen level (1%, for example), we judge the agreement to be unacceptable at that level. [For example, if $\text{Prob}(\text{outside } t\sigma)$ is less than 1%, the agreement is unacceptable at the 1% level.]

Problems for Chapter 5

For Section 5.1: Histograms and Distributions

5.1. ★ Draw a bar histogram for Joe's grades, as given in Quick Check 5.1. The horizontal axis should show the five possible grades (or, better, the corresponding scores $s_k = 0, 1, \dots, 4$) and the vertical axis, the fractions F_k . Compute his average grade as $\sum s_k F_k$.

5.2. ★ A health physicist places a weak radioactive sample in a liquid scintillation counter and records the number of decays in 10-minute intervals, starting at $t = 0$. His results are as follows (with t in minutes):

Interval for t :	0 to 10	10 to 20	20 to 30	30 to 40	40 to 50
Number:	9	6	3	1	1

and none beyond $t = 50$ min. Make a bin histogram of these events. Show t on the horizontal axis, and choose the vertical scale so that the area of each rectangle is the fraction of the decays that occurred in the corresponding bin.

5.3. ★★ A student measures the angular momenta L_i and L_f of a rotating system before and after adding an extra mass. To check the conservation of angular momentum, he calculates $L_i - L_f$ (expecting the answer zero). He repeats the measurement 50 times and collects his answers into bins as in Table 5.3, which shows his results

Table 5.3. Occurrences of values of $L_i - L_f$; for Problem 5.3.

After	Bin									
	(-9, -7)	(-7, -5)	(-5, -3)	(-3, -1)	(-1, 1)	(1, 3)	(3, 5)	(5, 7)	(7, 9)	
5 trials	0	1	2	0	1	0	1	0	0	
10 trials	0	1	2	2	3	1	1	0	0	
50 trials	1	3	7	8	10	9	6	4	2	

(in some unspecified units) after 5, 10, and 50 trials. Draw a bin histogram for each of these three cases. (Be careful to choose your scales so that the area of each rectangle is the fraction of events in the corresponding bin.)

5.4. ★★ A student makes 20 measurements of the time for a ball bearing to fall from the top to the bottom of a vertical cylinder of oil. She arranges her results in increasing order and counts how many times she got each different value, as follows (with the times in tenths of a second):

Time, t :	71	72	73	74	75	76	77	78	79	80
Occurrences:	2	0	3	5	4	1	3	1	0	1

(a) Draw a bin histogram of these results using bins of width 1, starting at 70.5. (Notice that, with bins of width 1, this bin histogram gives essentially the same

result as a bar histogram.) (b) Redraw the histogram with a bin width of 2, again starting at 70.5. (c) Notice how the wider bins in part (b) give a smoother histogram. If the chosen bins are too wide, however, information starts to be lost. To illustrate this loss, redraw the histogram with a bin width of 10, again starting at 70.5.

For Section 5.2: Limiting Distributions

5.5. ★ The limiting distribution for the results in some hypothetical measurement is given by the triangular function shown in Figure 5.18, where the value of $f(0)$ is

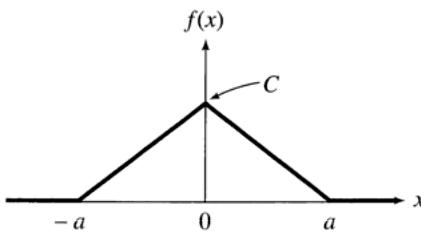


Figure 5.18. A triangular distribution; for Problem 5.5.

called C . (a) What is the probability of a measurement outside the range between $x = -a$ and $x = a$? (b) What is the probability of a measurement with $x > 0$? (c) Use the normalization condition (5.13) to find C in terms of a . (d) Sketch this function for the cases that $a = 1$ and $a = 2$.

5.6. ★ Consider an experiment similar to the one described in Problem 5.2, in which the experimenter counts the number of decays from a radioactive sample in a short time interval Δt from a radioactive source, starting at time $t = 0$. The limiting distribution for this kind of experiment is the *exponential distribution*,

$$f(t) = \frac{1}{\tau} e^{-t/\tau}, \quad (5.69)$$

where τ is a positive constant. (a) Sketch this function. (The distribution is zero for $t < 0$ because the experiment begins only at $t = 0$.) (b) Prove that this function satisfies the normalization condition (5.13). (c) Find the mean time \bar{t} at which the decays occur, as given by Equation (5.15). (Your answer here shows the significance of the parameter τ : It is the mean time at which the atoms in a large sample decay.)

5.7. ★ For the exponential distribution of Problem 5.6, what is the probability for a result $t > \tau$? What for $t > 2\tau$? (Notice that these probabilities also give the fraction of the original atoms that live longer than τ and 2τ .)

5.8. ★★ The physicist of Problem 5.2 decides that the limiting distribution for his experiment is the exponential distribution (5.69) (from Problem 5.6) with the parameter $\tau = 14.4$ min. If you have not already done so, draw the histogram for the data of Problem 5.2, and then draw the supposed limiting distribution on the same plot. How well does it seem to match? (You have no quantitative way to measure the agreement, so you can decide only if it *looks* satisfactory.)

5.9. ★★ The limiting distribution for the results in some hypothetical measurement has the form

$$f(x) = \begin{cases} C & \text{for } |x| < a \\ 0 & \text{otherwise.} \end{cases}$$

- (a) Use the normalization condition (5.13) to find C in terms of a . (b) Sketch this function. Describe in words the distribution of results governed by this distribution. (c) Use Equations (5.15) and (5.16) to calculate the mean and standard deviation that would be found after many measurements.

5.10. ★★ Explain clearly why the standard deviation for a limiting distribution $f(x)$ is given by (5.16). Your argument will parallel closely that leading from (5.14) to (5.15).

For Section 5.3: The Normal Distribution

5.11. ★ Using proper squared paper and clearly labeled axes, make good plots of the Gauss distribution

$$G_{X,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-X)^2/2\sigma^2}$$

for $X = 2$, $\sigma = 1$, and for $X = 3$, $\sigma = 0.3$. Use your calculator to compute the values of $G_{X,\sigma}(x)$. (If it has two memories to store $\sigma\sqrt{2\pi}$ and $-2\sigma^2$, this feature will speed your calculations. If you remember that the function is symmetric about $x = X$, the number of calculations needed is halved.) Put both graphs on the same plot for comparison.

5.12. ★ The width of a Gauss distribution is usually characterized by the parameter σ . An alternative parameter with a simple geometric interpretation is the *full width at half maximum*, or FWHM. This parameter is the distance between the two points x where $G_{X,\sigma}(x)$ is half its maximum value, as in Figure 5.19. Prove that

$$\text{FWHM} = 2\sigma\sqrt{2 \ln 2} = 2.35\sigma.$$

Some physicists use the *half width at half maximum*, shown in Figure 5.19 and defined as half the FWHM. Obviously HWHM = 1.17σ , or, very roughly, HWHM $\approx \sigma$.

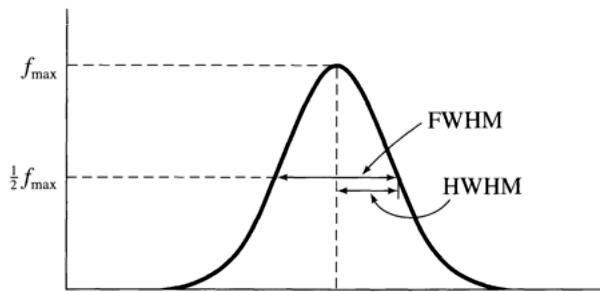


Figure 5.19. The full width at half maximum (FWHM) and the half width at half maximum (HWHM); for Problem 5.12.

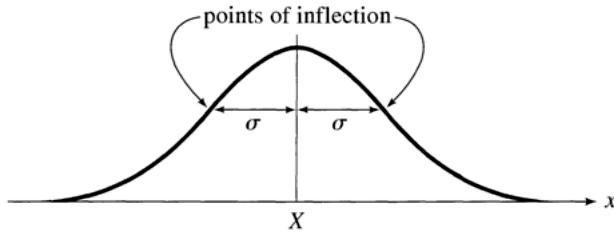


Figure 5.20. The points $X \pm \sigma$ are the points of inflection of the Gauss curve; for Problem 5.13.

5.13. ★ One way to define the width σ of the Gauss distribution is that the points $X \pm \sigma$ are the two points of inflection (Figure 5.20), where the curvature changes sign; that is, where the second derivative is zero. Prove this claim.

5.14. ★★ Make careful sketches of the normal distributions $G_{X,\sigma}(x)$ for the three cases:

- (a) $X = 0, \sigma = 1$
- (b) $X = 0, \sigma = 2$
- (c) $X = 5, \sigma = 1$

Describe in words the differences among the three curves.

5.15. ★★ If you have not yet done so, plot the third histogram of Problem 5.3. The student of that problem decides that the distribution of his results is consistent with the Gauss distribution $G_{X,\sigma}(x)$ centered on $X = 0$ with width $\sigma = 3.4$. Draw this distribution on the same graph and compare it with your histogram. (Read the hints to Problem 5.11. Note that you have no quantitative way to assess the fit; all you can do is see if the Gauss function *seems* to fit the histogram satisfactorily.)

5.16. ★★ Give in detail the steps leading from (5.30) to (5.31) to show that the standard deviation σ_x of many measurements normally distributed with width parameter σ is $\sigma_x = \sigma$.

For Section 5.4: The Standard Deviation as 68% Confidence Limit

5.17. ★ Some statisticians speak of a “68–95 rule.” What do you suppose this rule is? What is the “68–95–99.7 rule”?

5.18. ★ We often say that the probability of a measurement in the range $X \pm 2\sigma$ is about 95%. What is this probability to four significant figures? What range $X \pm t\sigma$ actually has probability equal to 95.00%? (Give t to three significant figures. The needed probabilities can be found in Appendix A.)

5.19. ★★ A student measures a quantity y many times and calculates his mean as $\bar{y} = 23$ and his standard deviation as $\sigma_y = 1$. What fraction of his readings would you expect to find between

- | | |
|--|--|
| <ul style="list-style-type: none"> (a) 22 and 24? (b) 22.5 and 23.5? (c) 21 and 25? | <ul style="list-style-type: none"> (d) 21 and 23? (e) 24 and 25? |
|--|--|

Finally, (f) within what limits (equidistant on either side of the mean) would you expect to find 50% of his readings? (The necessary information for all parts of this question is in Figure 5.13. More detailed information on these kinds of probabilities is in Appendixes A and B.)

5.20. ★★ An extensive survey reveals that the heights of men in a certain country are normally distributed, with a mean $\bar{h} = 69"$ and standard deviation $\sigma = 2"$. In a random sample of 1,000 men, how many would you expect to have a height

- (a) between 67" and 71"?
- (b) more than 71"?
- (c) more than 75"?
- (d) between 65" and 67"?

5.21. ★★ The Giraffe Club of Casterbridge is a club for young adults (18–24 years) who are unusually tall. There are 2,000 women between the ages of 18 and 24 in Casterbridge, and the heights of women in this age range are distributed normally with a mean of 5'5½" and standard deviation of 2½". (a) If the Giraffe Club initially sets its minimum height for women at 5'10", approximately how many women are eligible to join? (b) A year or so later, the club decides to double its female membership by lowering the minimum height requirement; what would you recommend that the club set as its new minimum height for women (to the nearest half inch)?

5.22. ★★ If the measurements of a quantity x are governed by the Gauss distribution $G_{X,\sigma}(x)$, the probability of obtaining a value between $X - t\sigma$ and $X + t\sigma$ is

$$\text{Prob}(\text{within } t\sigma) = \int_{X-t\sigma}^{X+t\sigma} G_{X,\sigma}(x) dx.$$

Prove carefully, showing all the necessary changes of variables, that

$$\text{Prob}(\text{within } t\sigma) = \frac{1}{\sqrt{2\pi}} \int_{-t}^t e^{-z^2/2} dz. \quad (5.70)$$

With each change of variables, check carefully what happens to your limits of integration. The integral (5.70) is often called the *error function*, denoted $\text{erf}(t)$, or the *normal error integral*.

5.23. ★★★ We have seen that for the normal distribution, the standard deviation gives the 68% confidence range. This result is not necessarily true for other distributions, as the following problem illustrates: Consider the exponential distribution of Problem 5.6, $f(t) = (1/\tau)e^{-t/\tau}$ (for $t \geq 0$; $f(t) = 0$ for $t < 0$). The parameter τ is the mean value of t (that is, after many measurements, $\bar{t} = \tau$). (a) Use the integral (5.16) to prove that τ is also the standard deviation, $\sigma_\tau = \tau$. (This result is a noteworthy property of this distribution—that the mean value is equal to the standard deviation.) (b) By doing the necessary integral, find the probability that any one value would fall in the range $\bar{t} \pm \sigma_\tau$.

For Section 5.5: Justification of the Mean as Best Estimate

5.24. ★ Suppose you have measured a quantity x six times, as follows:

$$51, 53, 54, 55, 52, 53.$$

(a) Assuming these measurements are normally distributed, what should be your best estimates for the true value X and the standard deviation σ ? (b) Based on these estimates, what is the probability that a seventh measurement would fall outside the range of the first six? (Given that your results are rounded to the nearest integer, this probability is that for a result $x \leq 50.5$ or $x \geq 55.5$.)

5.25. ★ A student measures a time t eight times with the following results (in tenths of a second):

Value, t_k :	75	76	77	78	79	80
Occurrences, n_k :	2	3	0	0	2	1

(a) Assuming these measurements are normally distributed, what should be your best estimates for the true value and the standard deviation? (b) Based on these estimates, what is the probability that a ninth measurement would be 81 or more? (Because the measurements are rounded to the nearest integer, this probability is that for a value $t \geq 80.5$.)

5.26. ★★ Suppose we have N measurements x_1, \dots, x_N of the same quantity x , and we believe that their limiting distribution should be the Gauss function $G_{X,\sigma}(x)$, with X and σ unknown. The principle of maximum likelihood asserts that the best estimate for the width is the value of σ for which the probability $Prob_{X,\sigma}(x_1, \dots, x_N)$ of the observed values x_1, \dots, x_N is largest. Differentiate $Prob_{X,\sigma}(x_1, \dots, x_N)$ in (5.40) with respect to σ , and show that the maximum occurs when σ is given by (5.43). [As discussed after (5.43), this result means that the best estimate for the true width σ is the standard deviation of the N observed values x_1, \dots, x_N . In practice, the true value X must be replaced with its best estimate \bar{x} , which requires replacement of N by $N - 1$, as proved in Appendix E.]

5.27. ★★ (a) Based on the data of Problem 5.24, find the mean, the standard deviation, and the uncertainty of your value for the SD [the last using Equation (5.46)]. (b) Recalculate the probability asked for in part (b) of Problem 5.24 assuming the true value of σ is $\sigma_x - \delta\sigma_x$, and, once again, assuming σ is really $\sigma_x + \delta\sigma_x$. Comment on the difference in your two answers here.

5.28. ★★ Based on several measurements of the same quantity x normally distributed about X with width σ , we can estimate X and σ . (a) Approximately how many measurements must we make to know σ within 30%? (b) Within 10%? (c) Within 3%?

Section 5.6: Justification of Addition in Quadrature

5.29. ★ The measurements of a certain quantity x are distributed according to the Gauss function $G_{X,\sigma}(x)$ with $X = 10$ and $\sigma = 2$. Sketch this distribution and the distributions for $q = x + 5$ and for $q' = x/2$, all on the same graph.

5.30. ★ Verify the identity (5.54) used in justifying addition in quadrature when propagating random errors.

For Section 5.7: Standard Deviation of the Mean

5.31. ★★ Listed here are 40 measurements t_1, \dots, t_{40} of the time for a stone to fall from a window to the ground (all in hundredths of a second).

63	58	74	78	70	74	75	82	68	69
76	62	72	88	65	81	79	77	66	76
86	72	79	77	60	70	65	69	73	77
72	79	65	66	70	74	84	76	80	69

(a) Compute the standard deviation σ_t for the 40 measurements. (b) Compute the means $\bar{t}_1, \dots, \bar{t}_{10}$ of the four measurements in each of the 10 columns. You can think of the data as resulting from 10 experiments, in each of which you found the *mean of four timings*. Given the result of part (a), what would you expect for the standard deviation of the 10 averages $\bar{t}_1, \dots, \bar{t}_{10}$? What is it? (c) Plot histograms for the 40 individual measurements t_1, \dots, t_{40} and for the 10 averages $\bar{t}_1, \dots, \bar{t}_{10}$. [Use the same scales and bin sizes for both plots so they can be compared easily. Bin boundaries can be chosen in various ways; perhaps the simplest is to put one boundary at the mean of all 40 measurements (72.90) and to use bins whose width is the standard deviation of the 10 averages $\bar{t}_1, \dots, \bar{t}_{10}$.]

5.32. ★★ In Problem 4.13 are listed 30 measurements of a certain time t . (a) If you haven't already done so, find the standard deviation of these 30 values. (b) Now think of the 30 data as 10 columns, each representing an experiment consisting of three measurements, and compute the 10 averages, $\bar{t}_1, \dots, \bar{t}_{10}$, for these 10 experiments. (c) Given the result of part (a), what would you expect for the standard deviation of the 10 means of part (b)? What is it? (d) Draw histograms for the 30 separate measurements and for the 10 means, $\bar{t}_1, \dots, \bar{t}_{10}$. (Use the same scale for both. A good way to choose your bin boundaries is to put one at the mean of all 30 measurements and to use the SD of the 10 means as the bin width.)

5.33. ★★ Based on the data of Problem 5.25, what would you give for the best estimate for the time t and its uncertainty? What is the uncertainty in your value for the uncertainty in t ? [The uncertainty in t is the SDOM, $\sigma_{\bar{t}} = \sigma_t/\sqrt{N}$. You can find the uncertainty in the SD σ_t from (5.46), and from this result you can find the uncertainty in the SDOM by error propagation.]

For Section 5.8: Acceptability of a Measured Answer

5.34. ★ According to a proposed theory, the quantity x should have the value x_{th} . Having measured x in the usual form $x_{\text{best}} \pm \sigma$ (where σ is the appropriate SD), we would say that the discrepancy between x_{best} and x_{th} is t standard deviations, where $t = |x_{\text{best}} - x_{\text{th}}|/\sigma$. How large must t be for us to say the discrepancy is significant at the 5% level? At the 2% level? At the 1% level?

5.35. ★ A student measures g , the acceleration of gravity, repeatedly and carefully and gets a final answer of 9.5 m/s^2 with a standard deviation of 0.1. If his measurements were normally distributed with center at the accepted value 9.8 and with width 0.1, what would be the probability of his getting an answer that differs from

9.8 by as much as (or more than) his? Assuming he made no actual mistakes, do you think his experiment may have suffered from some undetected systematic errors?

5.36. ★★ Two students measure the same quantity x and get final answers $x_A = 13 \pm 1$ and $x_B = 15 \pm 1$, where the quoted uncertainties are the appropriate standard deviations. **(a)** Assuming all errors are independent and random, what is the discrepancy $x_A - x_B$, and what is its uncertainty? **(b)** Assuming all quantities were normally distributed as expected, what would be the probability of getting a discrepancy as large as they did? Do you consider their discrepancy significant (at the 5% level)?

5.37. ★★ A nuclear physicist wants to check the conservation of energy in a certain nuclear reaction and measures the initial and final energies as $E_i = 75 \pm 3$ MeV and $E_f = 60 \pm 9$ MeV, where both quoted uncertainties are the standard deviations of the answers. Is this discrepancy significant (at the 5% level)? Explain your reasoning clearly.