3

Errors during the measurement process

3.1 Introduction

Errors in measurement systems can be divided into those that arise during the measurement process and those that arise due to later corruption of the measurement signal by induced noise during transfer of the signal from the point of measurement to some other point. This chapter considers only the first of these, with discussion on induced noise being deferred to Chapter 5.

It is extremely important in any measurement system to reduce errors to the minimum possible level and then to quantify the maximum remaining error that may exist in any instrument output reading. However, in many cases, there is a further complication that the final output from a measurement system is calculated by combining together two or more measurements of separate physical variables. In this case, special consideration must also be given to determining how the calculated error levels in each separate measurement should be combined together to give the best estimate of the most likely error magnitude in the calculated output quantity. This subject is considered in section 3.6.

The starting point in the quest to reduce the incidence of errors arising during the measurement process is to carry out a detailed analysis of all error sources in the system. Each of these error sources can then be considered in turn, looking for ways of eliminating or at least reducing the magnitude of errors. Errors arising during the measurement process can be divided into two groups, known as systematic errors and random errors.

Systematic errors describe errors in the output readings of a measurement system that are consistently on one side of the correct reading, i.e. either all the errors are positive or they are all negative. Two major sources of systematic errors are system disturbance during measurement and the effect of environmental changes (modifying inputs), as discussed in sections 3.2.1 and 3.2.2. Other sources of systematic error include bent meter needles, the use of uncalibrated instruments, drift in instrument characteristics and poor cabling practices. Even when systematic errors due to the above factors have been reduced or eliminated, some errors remain that are inherent in the manufacture of an instrument. These are quantified by the accuracy figure quoted in the published specifications contained in the instrument data sheet.

Random errors are perturbations of the measurement either side of the true value caused by random and unpredictable effects, such that positive errors and negative errors occur in approximately equal numbers for a series of measurements made of the same quantity. Such perturbations are mainly small, but large perturbations occur from time to time, again unpredictably. Random errors often arise when measurements are taken by human observation of an analogue meter, especially where this involves interpolation between scale points. Electrical noise can also be a source of random errors. To a large extent, random errors can be overcome by taking the same measurement a number of times and extracting a value by averaging or other statistical techniques, as discussed in section 3.5. However, any quantification of the measurement value and statement of error bounds remains a statistical quantity. Because of the nature of random errors and the fact that large perturbations in the measured quantity occur from time to time, the best that we can do is to express measurements in probabilistic terms: we may be able to assign a 95% or even 99% confidence level that the measurement is a certain value within error bounds of, say, ±1%, but we can never attach a 100% probability to measurement values that are subject to random errors.

Finally, a word must be said about the distinction between systematic and random errors. Error sources in the measurement system must be examined carefully to determine what type of error is present, systematic or random, and to apply the appropriate treatment. In the case of manual data measurements, a human observer may make a different observation at each attempt, but it is often reasonable to assume that the errors are random and that the mean of these readings is likely to be close to the correct value. However, this is only true as long as the human observer is not introducing a parallax-induced systematic error as well by persistently reading the position of a needle against the scale of an analogue meter from one side rather than from directly above. In that case, correction would have to be made for this systematic error (bias) in the measurements before statistical techniques were applied to reduce the effect of random errors.

3.2 Sources of systematic error

Systematic errors in the output of many instruments are due to factors inherent in the manufacture of the instrument arising out of tolerances in the components of the instrument. They can also arise due to wear in instrument components over a period of time. In other cases, systematic errors are introduced either by the effect of environmental disturbances or through the disturbance of the measured system by the act of measurement. These various sources of systematic error, and ways in which the magnitude of the errors can be reduced, are discussed below.

3.2.1 System disturbance due to measurement

Disturbance of the measured system by the act of measurement is a common source of systematic error. If we were to start with a beaker of hot water and wished to measure its temperature with a mercury-in-glass thermometer, then we would take the

thermometer, which would initially be at room temperature, and plunge it into the water. In so doing, we would be introducing a relatively cold mass (the thermometer) into the hot water and a heat transfer would take place between the water and the thermometer. This heat transfer would lower the temperature of the water. Whilst the reduction in temperature in this case would be so small as to be undetectable by the limited measurement resolution of such a thermometer, the effect is finite and clearly establishes the principle that, in nearly all measurement situations, the process of measurement disturbs the system and alters the values of the physical quantities being measured.

A particularly important example of this occurs with the orifice plate. This is placed into a fluid-carrying pipe to measure the flow rate, which is a function of the pressure that is measured either side of the orifice plate. This measurement procedure causes a permanent pressure loss in the flowing fluid. The disturbance of the measured system can often be very significant.

Thus, as a general rule, the process of measurement always disturbs the system being measured. The magnitude of the disturbance varies from one measurement system to the next and is affected particularly by the type of instrument used for measurement. Ways of minimizing disturbance of measured systems is an important consideration in instrument design. However, an accurate understanding of the mechanisms of system disturbance is a prerequisite for this.

Measurements in electric circuits

In analysing system disturbance during measurements in electric circuits, Thévenin's theorem (see Appendix 3) is often of great assistance. For instance, consider the circuit shown in Figure 3.1(a) in which the voltage across resistor R_5 is to be measured by a voltmeter with resistance R_m. Here, R_m acts as a shunt resistance across R_5, decreasing the resistance between points AB and so disturbing the circuit. Therefore, the voltage E_m measured by the meter is not the value of the voltage E_0 that existed prior to measurement. The extent of the disturbance can be assessed by calculating the open-circuit voltage E_0 and comparing it with E_m.

Thévenin's theorem allows the circuit of Figure 3.1(a) comprising two voltage sources and five resistors to be replaced by an equivalent circuit containing a single resistance and one voltage source, as shown in Figure 3.1(b). For the purpose of defining the equivalent single resistance of a circuit by Thévenin's theorem, all voltage sources are represented just by their internal resistance, which can be approximated to zero, as shown in Figure 3.1(c). Analysis proceeds by calculating the equivalent resistances of sections of the circuit and building these up until the required equivalent resistance of the whole of the circuit is obtained. Starting at C and D, the circuit to the left of C and D consists of a series pair of resistances (R_1 and R_2) in parallel with R_3, and the equivalent resistance can be written as:

$$\frac{1}{R_{CD}} = \frac{1}{R_1 + R_2} + \frac{1}{R_3} \quad \text{or} \quad R_{CD} = \frac{(R_1 + R_2)R_3}{R_1 + R_2 + R_3}$$

Moving now to A and B, the circuit to the left consists of a pair of series resistances (R_{CD} and R_4) in parallel with R_5. The equivalent circuit resistance R_{AB} can thus be

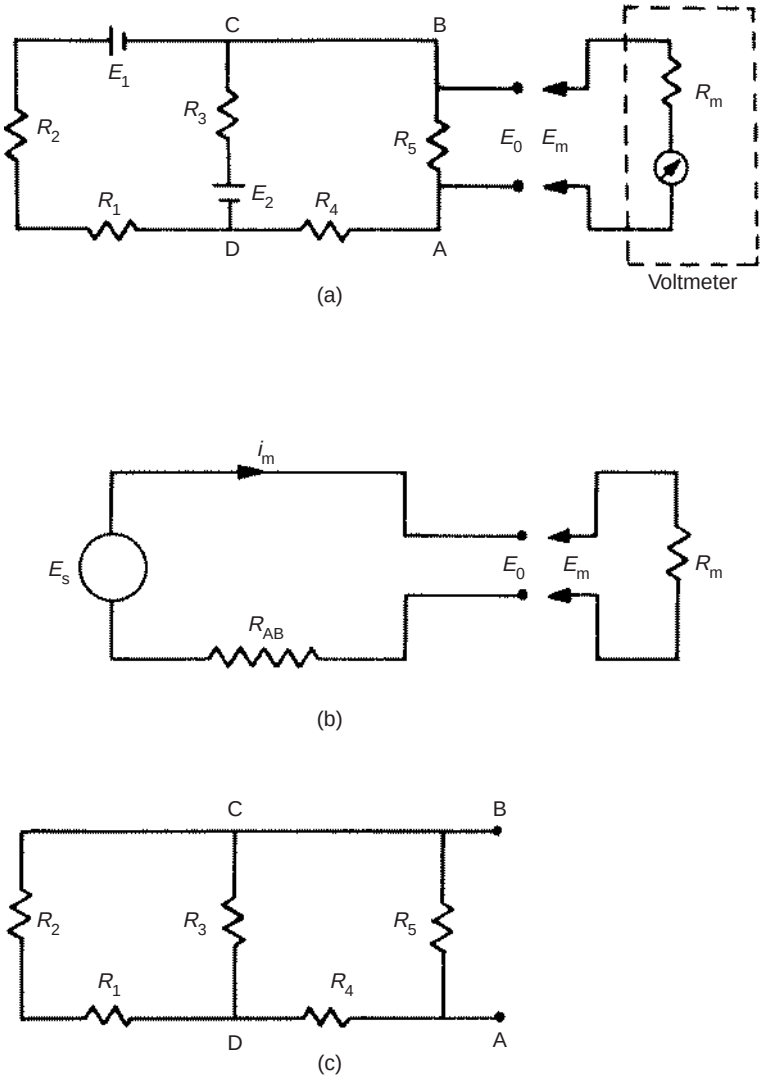

Fig. 3.1 Analysis of circuit loading: (a) a circuit in which the voltage across R_5 is to be measured; (b) equivalent circuit by Thévenin's theorem; (c) the circuit used to find the equivalent single resistance R_{AB}.

written as:

$$\frac{1}{R_{AB}} = \frac{1}{R_{CD} + R_4} + \frac{1}{R_5} \quad \text{or} \quad R_{AB} = \frac{(R_4 + R_{CD})R_5}{R_4 + R_{CD} + R_5}$$

Substituting for R_{CD} using the expression derived previously, we obtain:

$$R_{AB} = \frac{\left[\dfrac{(R_1 + R_2)R_3}{R_1 + R_2 + R_3} + R_4\right] R_5}{\dfrac{(R_1 + R_2)R_3}{R_1 + R_2 + R_3} + R_4 + R_5} \tag{3.1}$$

Defining I as the current flowing in the circuit when the measuring instrument is connected to it, we can write:

$$I = \frac{E_0}{R_{AB} + R_m},$$

and the voltage measured by the meter is then given by:

$$E_m = \frac{R_m E_0}{R_{AB} + R_m}.$$

In the absence of the measuring instrument and its resistance R_m, the voltage across AB would be the equivalent circuit voltage source whose value is E_0. The effect of measurement is therefore to reduce the voltage across AB by the ratio given by:

$$\frac{E_m}{E_0} = \frac{R_m}{R_{AB} + R_m} \tag{3.2}$$

It is thus obvious that as R_m gets larger, the ratio E_m/E_0 gets closer to unity, showing that the design strategy should be to make R_m as high as possible to minimize disturbance of the measured system. (Note that we did not calculate the value of E_0, since this is not required in quantifying the effect of R_m.)

Example 3.1
Suppose that the components of the circuit shown in Figure 3.1(a) have the following values:

$$R_1 = 400\,\Omega; R_2 = 600\,\Omega; R_3 = 1000\,\Omega, R_4 = 500\,\Omega; R_5 = 1000\,\Omega$$

The voltage across AB is measured by a voltmeter whose internal resistance is 9500 Ω. What is the measurement error caused by the resistance of the measuring instrument?

Solution
Proceeding by applying Thévenin's theorem to find an equivalent circuit to that of Figure 3.1(a) of the form shown in Figure 3.1(b), and substituting the given component values into the equation for R_{AB} (3.1), we obtain:

$$R_{AB} = \frac{[(1000^2/2000) + 500]1000}{(1000^2/2000) + 500 + 1000} = \frac{1000^2}{2000} = 500\,\Omega$$

From equation (3.2), we have:

$$\frac{E_m}{E_0} = \frac{R_m}{R_{AB} + R_m}$$

The measurement error is given by $(E_0 - E_m)$:

$$E_0 - E_m = E_0\left(1 - \frac{R_m}{R_{AB} + R_m}\right)$$

Substituting in values:

$$E_0 - E_m = E_0\left(1 - \frac{9500}{10\,000}\right) = 0.95E_0$$

Thus, the error in the measured value is 5%.

At this point, it is interesting to note the constraints that exist when practical attempts are made to achieve a high internal resistance in the design of a moving-coil voltmeter. Such an instrument consists of a coil carrying a pointer mounted in a fixed magnetic field. As current flows through the coil, the interaction between the field generated and the fixed field causes the pointer it carries to turn in proportion to the applied current (see Chapter 6 for further details). The simplest way of increasing the input impedance (the resistance) of the meter is either to increase the number of turns in the coil or to construct the same number of coil turns with a higher-resistance material. However, either of these solutions decreases the current flowing in the coil, giving less magnetic torque and thus decreasing the measurement sensitivity of the instrument (i.e. for a given applied voltage, we get less deflection of the pointer). This problem can be overcome by changing the spring constant of the restraining springs of the instrument, such that less torque is required to turn the pointer by a given amount. However, this reduces the ruggedness of the instrument and also demands better pivot design to reduce friction. This highlights a very important but tiresome principle in instrument design: any attempt to improve the performance of an instrument in one respect generally decreases the performance in some other aspect. This is an inescapable fact of life with passive instruments such as the type of voltmeter mentioned, and is often the reason for the use of alternative active instruments such as digital voltmeters, where the inclusion of auxiliary power greatly improves performance.

Bridge circuits for measuring resistance values are a further example of the need for careful design of the measurement system. The impedance of the instrument measuring the bridge output voltage must be very large in comparison with the component resistances in the bridge circuit. Otherwise, the measuring instrument will load the circuit and draw current from it. This is discussed more fully in Chapter 7.

3.2.2 Errors due to environmental inputs

An environmental input is defined as an apparently real input to a measurement system that is actually caused by a change in the environmental conditions surrounding the measurement system. The fact that the static and dynamic characteristics specified for measuring instruments are only valid for particular environmental conditions (e.g. of temperature and pressure) has already been discussed at considerable length in Chapter 2. These specified conditions must be reproduced as closely as possible during calibration exercises because, away from the specified calibration conditions, the characteristics of measuring instruments vary to some extent and cause measurement errors. The magnitude of this environment-induced variation is quantified by the two constants known as sensitivity drift and zero drift, both of which are generally included in the published specifications for an instrument. Such variations of environmental conditions away from the calibration conditions are sometimes described as modifying inputs to the measurement system because they modify the output of the system. When such modifying inputs are present, it is often difficult to determine how much of the output change in a measurement system is due to a change in the measured variable and how much is due to a change in environmental conditions. This is illustrated by the following example. Suppose we are given a small closed box and told that it may contain either a mouse or a rat. We are also told that the box weighs 0.1 kg when

empty. If we put the box onto bathroom scales and observe a reading of 1.0 kg, this does not immediately tell us what is in the box because the reading may be due to one of three things:

(a) a 0.9 kg rat in the box (real input)
(b) an empty box with a 0.9 kg bias on the scales due to a temperature change (environmental input)
(c) a 0.4 kg mouse in the box together with a 0.5 kg bias (real + environmental inputs).

Thus, the magnitude of any environmental input must be measured before the value of the measured quantity (the real input) can be determined from the output reading of an instrument.

In any general measurement situation, it is very difficult to avoid environmental inputs, because it is either impractical or impossible to control the environmental conditions surrounding the measurement system. System designers are therefore charged with the task of either reducing the susceptibility of measuring instruments to environmental inputs or, alternatively, quantifying the effect of environmental inputs and correcting for them in the instrument output reading. The techniques used to deal with environmental inputs and minimize their effect on the final output measurement follow a number of routes as discussed below.

3.2.3 Wear in instrument components

Systematic errors can frequently develop over a period of time because of wear in instrument components. Recalibration often provides a full solution to this problem.

3.2.4 Connecting leads

In connecting together the components of a measurement system, a common source of error is the failure to take proper account of the resistance of connecting leads (or pipes in the case of pneumatically or hydraulically actuated measurement systems). For instance, in typical applications of a resistance thermometer, it is common to find that the thermometer is separated from other parts of the measurement system by perhaps 100 metres. The resistance of such a length of 20 gauge copper wire is 7 Ω, and there is a further complication that such wire has a temperature coefficient of 1 mΩ/°C.

Therefore, careful consideration needs to be given to the choice of connecting leads. Not only should they be of adequate cross-section so that their resistance is minimized, but they should be adequately screened if they are thought likely to be subject to electrical or magnetic fields that could otherwise cause induced noise. Where screening is thought essential, then the routing of cables also needs careful planning. In one application in the author's personal experience involving instrumentation of an electric-arc steel making furnace, screened signal-carrying cables between transducers on the arc furnace and a control room at the side of the furnace were initially corrupted by high amplitude 50 Hz noise. However, by changing the route of the cables between the transducers and the control room, the magnitude of this induced noise was reduced by a factor of about ten.

3.3 Reduction of systematic errors

The prerequisite for the reduction of systematic errors is a complete analysis of the measurement system that identifies all sources of error. Simple faults within a system, such as bent meter needles and poor cabling practices, can usually be readily and cheaply rectified once they have been identified. However, other error sources require more detailed analysis and treatment. Various approaches to error reduction are considered below.

3.3.1 Careful instrument design

Careful instrument design is the most useful weapon in the battle against environmental inputs, by reducing the sensitivity of an instrument to environmental inputs to as low a level as possible. For instance, in the design of strain gauges, the element should be constructed from a material whose resistance has a very low temperature coefficient (i.e. the variation of the resistance with temperature is very small). However, errors due to the way in which an instrument is designed are not always easy to correct, and a choice often has to be made between the high cost of redesign and the alternative of accepting the reduced measurement accuracy if redesign is not undertaken.

3.3.2 Method of opposing inputs

The method of opposing inputs compensates for the effect of an environmental input in a measurement system by introducing an equal and opposite environmental input that cancels it out. One example of how this technique is applied is in the type of millivoltmeter shown in Figure 3.2. This consists of a coil suspended in a fixed magnetic field produced by a permanent magnet. When an unknown voltage is applied to the coil, the magnetic field due to the current interacts with the fixed field and causes the coil (and a pointer attached to the coil) to turn. If the coil resistance R_{coil} is sensitive to temperature, then any environmental input to the system in the form of a temperature change will alter the value of the coil current for a given applied voltage and so alter the pointer output reading. Compensation for this is made by introducing a compensating resistance R_{comp} into the circuit, where R_{comp} has a temperature coefficient that is equal in magnitude but opposite in sign to that of the coil. Thus, in response to an increase in temperature, R_{coil} increases but R_{comp} decreases, and so the total resistance remains approximately the same.

3.3.3 High-gain feedback

The benefit of adding high-gain feedback to many measurement systems is illustrated by considering the case of the voltage-measuring instrument whose block diagram is shown in Figure 3.3. In this system, the unknown voltage E_i is applied to a motor of torque constant K_{m}, and the induced torque turns a pointer against the restraining action of a spring with spring constant K_{s}. The effect of environmental inputs on the

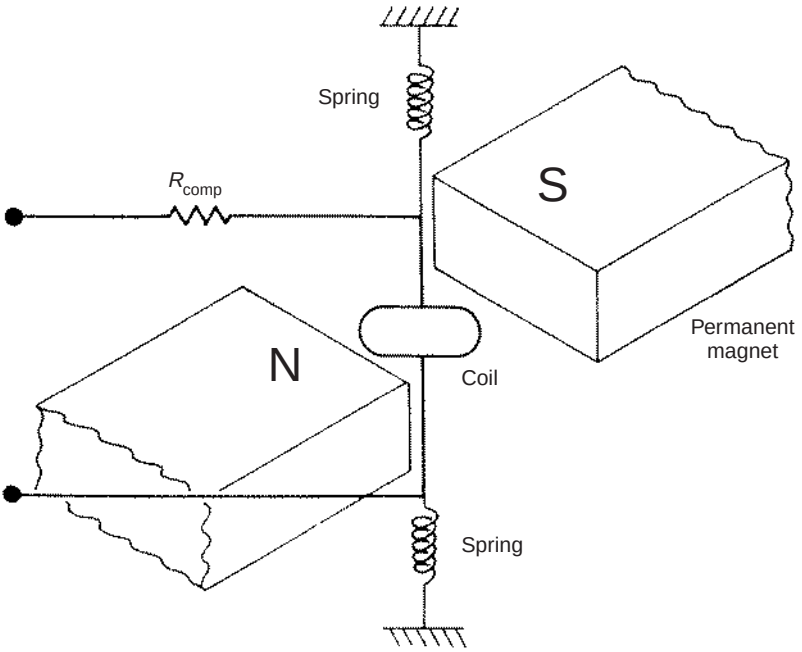

Fig. 3.2 Millivoltmeter.

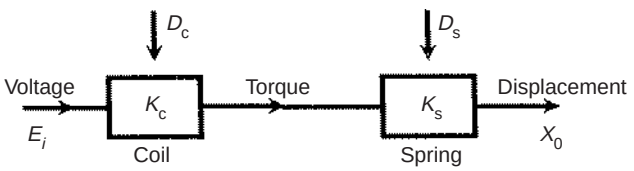

Fig. 3.3 Block diagram for voltage-measuring instrument.

motor and spring constants is represented by variables D_m and D_s. In the absence of environmental inputs, the displacement of the pointer X_0 is given by: $X_0 = K_m K_s E_i$. However, in the presence of environmental inputs, both K_m and K_s change, and the relationship between X_0 and E_i can be affected greatly. Therefore, it becomes difficult or impossible to calculate E_i from the measured value of X_0. Consider now what happens if the system is converted into a high-gain, closed-loop one, as shown in Figure 3.4, by adding an amplifier of gain constant K_a and a feedback device with gain constant K_f. Assume also that the effect of environmental inputs on the values of K_a and K_f are represented by D_a and D_f. The feedback device feeds back a voltage E_0 proportional to the pointer displacement X_0. This is compared with the unknown voltage E_i by a comparator and the error is amplified. Writing down the equations of the system, we have:

$$E_0 = K_f X_0; \quad X_0 = (E_i - E_0) K_a K_m K_s = (E_i - K_f X_0) K_a K_m K_s$$

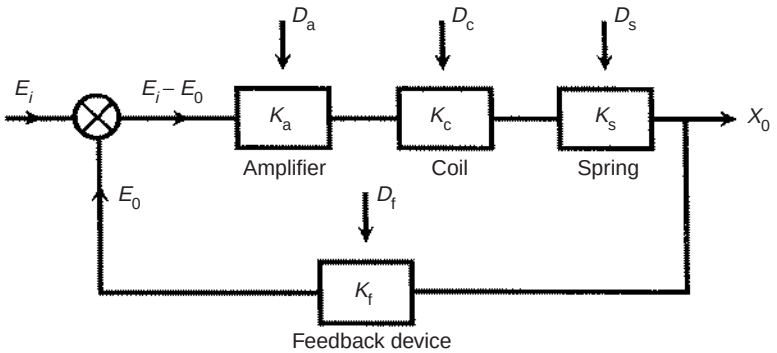

Fig. 3.4 Block diagram of voltage-measuring instrument with high-gain feedback.

Thus:

$$E_i K_a K_m K_s = (1 + K_f K_a K_m K_s) X_0$$

i.e.

$$X_0 = \frac{K_a K_m K_s}{1 + K_f K_a K_m K_s} E_i \tag{3.3}$$

Because K_a is very large (it is a high-gain amplifier), $K_f . K_a . K_m . K_s \gg 1$, and equation (3.3) reduces to:

$$X_0 = E_i / K_f$$

This is a highly important result because we have reduced the relationship between X_0 and E_i to one that involves only K_f. The sensitivity of the gain constants K_a, K_m and K_s to the environmental inputs D_a, D_m and D_s has thereby been rendered irrelevant, and we only have to be concerned with one environmental input D_f. Conveniently, it is usually easy to design a feedback device that is insensitive to environmental inputs: this is much easier than trying to make a motor or spring insensitive. Thus, high-gain feedback techniques are often a very effective way of reducing a measurement system's sensitivity to environmental inputs. However, one potential problem that must be mentioned is that there is a possibility that high-gain feedback will cause instability in the system. Therefore, any application of this method must include careful stability analysis of the system.

3.3.4 Calibration

Instrument calibration is a very important consideration in measurement systems and calibration procedures are considered in detail in Chapter 4. All instruments suffer drift in their characteristics, and the rate at which this happens depends on many factors, such as the environmental conditions in which instruments are used and the frequency of their use. Thus, errors due to instruments being out of calibration can usually be rectified by increasing the frequency of recalibration.

3.3.5 Manual correction of output reading

In the case of errors that are due either to system disturbance during the act of measurement or due to environmental changes, a good measurement technician can substantially reduce errors at the output of a measurement system by calculating the effect of such systematic errors and making appropriate correction to the instrument readings. This is not necessarily an easy task, and requires all disturbances in the measurement system to be quantified. This procedure is carried out automatically by intelligent instruments.

3.3.6 Intelligent instruments

Intelligent instruments contain extra sensors that measure the value of environmental inputs and automatically compensate the value of the output reading. They have the ability to deal very effectively with systematic errors in measurement systems, and errors can be attenuated to very low levels in many cases. A more detailed analysis of intelligent instruments can be found in Chapter 9.

3.4 Quantification of systematic errors

Once all practical steps have been taken to eliminate or reduce the magnitude of systematic errors, the final action required is to estimate the maximum remaining error that may exist in a measurement due to systematic errors. Unfortunately, it is not always possible to quantify exact values of a systematic error, particularly if measurements are subject to unpredictable environmental conditions. The usual course of action is to assume mid-point environmental conditions and specify the maximum measurement error as $\pm x\%$ of the output reading to allow for the maximum expected deviation in environmental conditions away from this mid-point. Data sheets supplied by instrument manufacturers usually quantify systematic errors in this way, and such figures take account of all systematic errors that may be present in output readings from the instrument.

3.5 Random errors

Random errors in measurements are caused by unpredictable variations in the measurement system. They are usually observed as small perturbations of the measurement either side of the correct value, i.e. positive errors and negative errors occur in approximately equal numbers for a series of measurements made of the same constant quantity. Therefore, random errors can largely be eliminated by calculating the average of a number of repeated measurements, provided that the measured quantity remains constant during the process of taking the repeated measurements. This averaging process of repeated measurements can be done automatically by intelligent instruments, as discussed in Chapter 9. The degree of confidence in the calculated mean/median values can be quantified by calculating the standard deviation or variance of the data,

these being parameters that describe how the measurements are distributed about the mean value/median. All of these terms are explained more fully in section 3.5.1.

3.5.1 Statistical analysis of measurements subject to random errors

Mean and median values

The average value of a set of measurements of a constant quantity can be expressed as either the mean value or the median value. As the number of measurements increases, the difference between the mean value and median values becomes very small. However, for any set of n measurements $x_1, x_2 \cdots x_n$ of a constant quantity, the most likely true value is the *mean* given by:

$$x_{\text{mean}} = \frac{x_1 + x_2 + \cdots x_n}{n} \tag{3.4}$$

This is valid for all data sets where the measurement errors are distributed equally about the zero error value, i.e. where the positive errors are balanced in quantity and magnitude by the negative errors.

The *median* is an approximation to the mean that can be written down without having to sum the measurements. The median is the middle value when the measurements in the data set are written down in ascending order of magnitude. For a set of n measurements $x_1, x_2 \cdots x_n$ of a constant quantity, written down in ascending order of magnitude, the median value is given by:

$$x_{\text{median}} = x_{n+1}/2 \tag{3.5}$$

Thus, for a set of 9 measurements $x_1, x_2 \cdots x_9$ arranged in order of magnitude, the median value is x_5. For an even number of measurements, the median value is midway between the two centre values, i.e. for 10 measurements $x_1 \cdots x_{10}$, the median value is given by: $(x_5 + x_6)/2$.

Suppose that the length of a steel bar is measured by a number of different observers and the following set of 11 measurements are recorded (units mm). We will call this measurement set A.

398 420 394 416 404 408 400 420 396 413 430 (Measurement set A)

Using (3.4) and (3.5), mean = 409.0 and median = 408. Suppose now that the measurements are taken again using a better measuring rule, and with the observers taking more care, to produce the following measurement set B:

409 406 402 407 405 404 407 404 407 407 408 (Measurement set B)

For these measurements, mean = 406.0 and median = 407. Which of the two measurement sets A and B, and the corresponding mean and median values, should we have most confidence in? Intuitively, we can regard measurement set B as being more reliable since the measurements are much closer together. In set A, the spread between the smallest (396) and largest (430) value is 34, whilst in set B, the spread is only 6.

- *Thus, the smaller the spread of the measurements, the more confidence we have in the mean or median value calculated.*

Let us now see what happens if we increase the number of measurements by extending measurement set B to 23 measurements. We will call this measurement set C.

409 406 402 407 405 404 407 404 407 407 408 406 410 406 405 408

406 409 406 405 409 406 407 (Measurement set C)

Now, mean = 406.5 and median = 406.

- *This confirms our earlier statement that the median value tends towards the mean value as the number of measurements increases.*

Standard deviation and variance

Expressing the spread of measurements simply as the range between the largest and smallest value is not in fact a very good way of examining how the measurement values are distributed about the mean value. A much better way of expressing the distribution is to calculate the variance or standard deviation of the measurements. The starting point for calculating these parameters is to calculate the deviation (error) d_i of each measurement x_i from the mean value x_{mean}:

$$d_i = x_i - x_{\text{mean}} \tag{3.6}$$

The *variance* (V) is then given by:*

$$V = \frac{d_1^2 + d_2^2 \cdots d_n^2}{n-1} \tag{3.7}$$

The *standard deviation* (σ) is simply the square root of the variance. Thus*:

$$\sigma = \sqrt{V} = \sqrt{\frac{d_1^2 + d_2^2 \cdots d_n^2}{n-1}} \tag{3.8}$$

Example 3.2
Calculate σ and V for measurement sets A, B and C above.

* Mathematically minded readers may have observed that the expressions for V and σ differ from the formal mathematical definitions, which have (n) instead of $(n-1)$ in the denominator. This difference arises because the mathematical definition is for an infinite data set, whereas, in the case of measurements, we are concerned only with finite data sets. For a finite set of measurements (x_i) $i = 1, n$, the mean x_{m} will differ from the true mean μ of the infinite data set that the finite set is part of. If somehow we knew the true mean μ of a set of measurements, then the deviations d_i could be calculated as the deviation of each data value from the true mean and it would then be correct to calculate V and σ using (n) instead of $(n-1)$. However, in normal situations, using $(n-1)$ in the denominator of equations (3.7) and (3.8) produces a value that is statistically closer to the correct value.

Solution
First, draw a table of measurements and deviations for set A (mean = 409 as calculated earlier):

Measurement	398	420	394	416	404	408	400	420	396	413	430
Deviation from mean	−11	+11	−15	+7	−5	−1	−9	+11	−13	+4	+21
(deviations)^2	121	121	225	49	25	1	81	121	169	16	441

$\sum(\text{deviations}^2) = 1370$; n = number of measurements = 11.
Then, from (3.7), $V = \sum(\text{deviations}^2)/n - 1; = 1370/10 = 137$; $\sigma = \sqrt{V} = 11.7$.
The measurements and deviations for set B are (mean = 406 as calculated earlier):

Measurement	409	406	402	407	405	404	407	404	407	407	408
Deviation from mean	+3	0	−4	+1	−1	−2	+1	−2	+1	+1	+2
(deviations)^2	9	0	16	1	1	4	1	4	1	1	4

From this data, using (3.7) and (3.8), $V = 4.2$ and $\sigma = 2.05$.
The measurements and deviations for set C are (mean = 406.5 as calculated earlier):

Measurement	409	406	402	407	405	404	407	404
Deviation from mean	+2.5	−0.5	−4.5	+0.5	−1.5	−2.5	+0.5	−2.5
(deviations)^2	6.25	0.25	20.25	0.25	2.25	6.25	0.25	6.25

Measurement	407	407	408	406	410	406	405	408
Deviation from mean	+0.5	+0.5	+1.5	−0.5	+3.5	−0.5	−1.5	+1.5
(deviations)^2	0.25	0.25	2.25	0.25	12.25	0.25	2.25	2.25

Measurement	406	409	406	405	409	406	407
Deviation from mean	−0.5	+2.5	−0.5	−1.5	+2.5	−0.5	+0.5
(deviations)^2	0.25	6.25	0.25	2.25	6.25	0.25	0.25

From this data, using (3.7) and (3.8), $V = 3.53$ and $\sigma = 1.88$.

Note that the smaller values of V and σ for measurement set B compared with A correspond with the respective size of the spread in the range between maximum and minimum values for the two sets.

- *Thus, as V and σ decrease for a measurement set, we are able to express greater confidence that the calculated mean or median value is close to the true value, i.e. that the averaging process has reduced the random error value close to zero.*
- *Comparing V and σ for measurement sets B and C, V and σ get smaller as the number of measurements increases, confirming that confidence in the mean value increases as the number of measurements increases.*

We have observed so far that random errors can be reduced by taking the average (mean or median) of a number of measurements. However, although the mean or median value is close to the true value, it would only become exactly equal to the true value if we could average an infinite number of measurements. As we can only make a finite number of measurements in a practical situation, the average value will still have some error. This error can be quantified as the *standard error of the mean*, which will be discussed in detail a little later. However, before that, the subject of graphical analysis of random measurement errors needs to be covered.

3.5.2 Graphical data analysis techniques – frequency distributions

Graphical techniques are a very useful way of analysing the way in which random measurement errors are distributed. The simplest way of doing this is to draw a *histogram*, in which bands of equal width across the range of measurement values are defined and the number of measurements within each band is counted. Figure 3.5 shows a histogram for set C of the length measurement data given in section 3.5.1, in which the bands chosen are 2 mm wide. For instance, there are 11 measurements in the range between 405.5 and 407.5 and so the height of the histogram for this range is 11 units. Also, there are 5 measurements in the range from 407.5 to 409.5 and so the height of the histogram over this range is 5 units. The rest of the histogram is completed in a similar fashion. (N.B. The scaling of the bands was deliberately chosen so that no measurements fell on the boundary between different bands and caused ambiguity about which band to put them in.) Such a histogram has the characteristic shape shown by truly random data, with symmetry about the mean value of the measurements.

As it is the actual value of measurement error that is usually of most concern, it is often more useful to draw a histogram of the deviations of the measurements

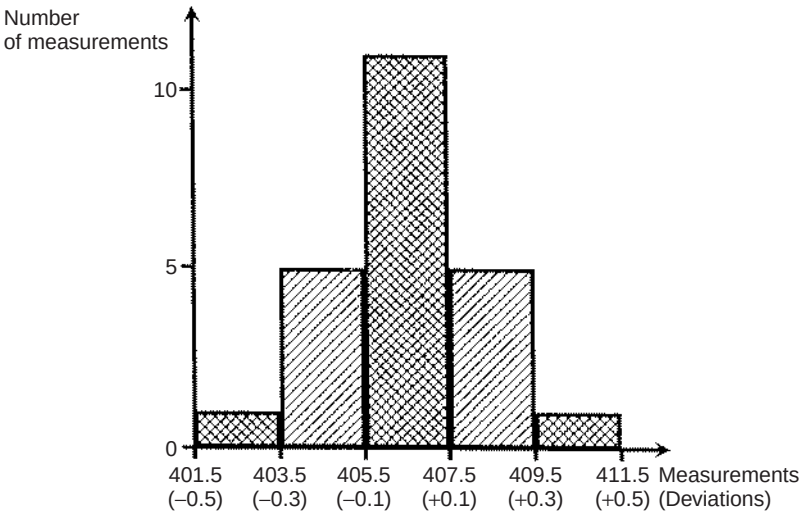

Fig. 3.5 Histogram of measurements and deviations.

from the mean value rather than to draw a histogram of the measurements themselves. The starting point for this is to calculate the deviation of each measurement away from the calculated mean value. Then a *histogram of deviations* can be drawn by defining deviation bands of equal width and counting the number of deviation values in each band. This histogram has exactly the same shape as the histogram of the raw measurements except that the scaling of the horizontal axis has to be redefined in terms of the deviation values (these units are shown in brackets on Figure 3.5).

Let us now explore what happens to the histogram of deviations as the number of measurements increases. As the number of measurements increases, smaller bands can be defined for the histogram, which retains its basic shape but then consists of a larger number of smaller steps on each side of the peak. In the limit, as the number of measurements approaches infinity, the histogram becomes a smooth curve known as a *frequency distribution curve* as shown in Figure 3.6. The ordinate of this curve is the frequency of occurrence of each deviation value, $F(D)$, and the abscissa is the magnitude of deviation, D.

The symmetry of Figures 3.5 and 3.6 about the zero deviation value is very useful for showing graphically that the measurement data only has random errors. Although these figures cannot easily be used to quantify the magnitude and distribution of the errors, very similar graphical techniques do achieve this. If the height of the frequency distribution curve is normalized such that the area under it is unity, then the curve in this form is known as a *probability curve*, and the height $F(D)$ at any particular deviation magnitude D is known as the *probability density function* (p.d.f.). The condition that

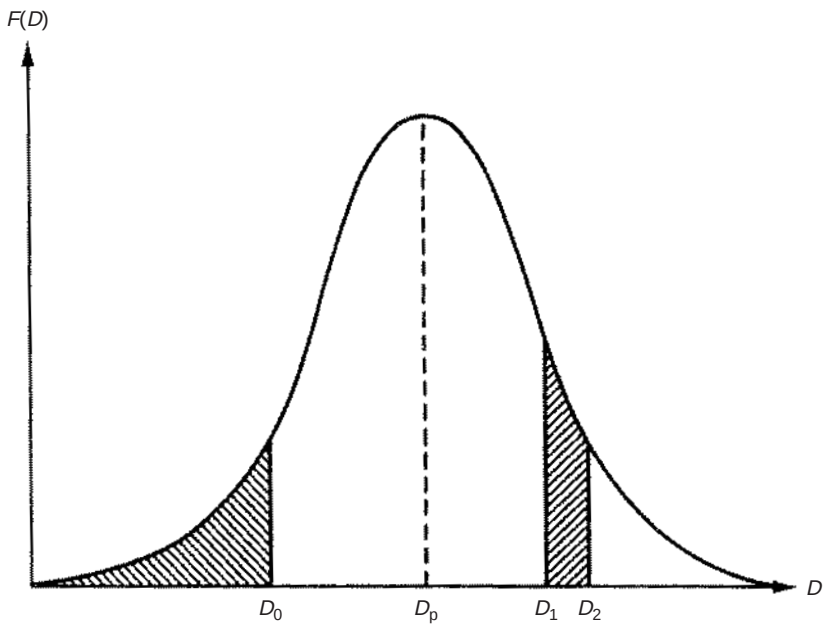

Fig. 3.6 Frequency distribution curve of deviations.

the area under the curve is unity can be expressed mathematically as:

$$\int_{-\infty}^{\infty} F(D)\,\mathrm{d}D = 1$$

The probability that the error in any one particular measurement lies between two levels D_1 and D_2 can be calculated by measuring the area under the curve contained between two vertical lines drawn through D_1 and D_2, as shown by the right-hand hatched area in Figure 3.6. This can be expressed mathematically as:

$$P(D_1 \leq D \leq D_2) = \int_{D_1}^{D_2} F(D)\,\mathrm{d}D \tag{3.9}$$

Of particular importance for assessing the maximum error likely in any one measurement is the *cumulative distribution function* (c.d.f.). This is defined as the probability of observing a value less than or equal to D_0, and is expressed mathematically as:

$$P(D \leq D_0) = \int_{-\infty}^{D_0} F(D)\,\mathrm{d}D \tag{3.10}$$

Thus, the c.d.f. is the area under the curve to the left of a vertical line drawn through D_0, as shown by the left-hand hatched area on Figure 3.6.

The deviation magnitude D_p corresponding with the peak of the frequency distribution curve (Figure 3.6) is the value of deviation that has the greatest probability. If the errors are entirely random in nature, then the value of D_p will equal zero. Any non-zero value of D_p indicates systematic errors in the data, in the form of a bias that is often removable by recalibration.

Gaussian distribution

Measurement sets that only contain random errors usually conform to a distribution with a particular shape that is called *Gaussian*, although this conformance must always be tested (see the later section headed 'Goodness of fit'). The shape of a Gaussian curve is such that the frequency of small deviations from the mean value is much greater than the frequency of large deviations. This coincides with the usual expectation in measurements subject to random errors that the number of measurements with a small error is much larger than the number of measurements with a large error. Alternative names for the Gaussian distribution are the *Normal distribution* or *Bell-shaped distribution*. A Gaussian curve is formally defined as a normalized frequency distribution that is symmetrical about the line of zero error and in which the frequency and magnitude of quantities are related by the expression:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{[-(x-m)^2/2\sigma^2]} \tag{3.11}$$

where m is the mean value of the data set x and the other quantities are as defined before. Equation (3.11) is particularly useful for analysing a Gaussian set of measurements and predicting how many measurements lie within some particular defined range. If the measurement deviations D are calculated for all measurements such that $D = x - m$, then the curve of deviation frequency $F(D)$ plotted against deviation magnitude D is

a Gaussian curve known as the *error frequency distribution curve*. The mathematical relationship between $F(D)$ and D can then be derived by modifying equation (3.11) to give:

$$F(D) = \frac{1}{\sigma\sqrt{2\pi}} e^{[-D^2/2\sigma^2]} \tag{3.12}$$

The shape of a Gaussian curve is strongly influenced by the value of σ, with the width of the curve decreasing as σ becomes smaller. As a smaller σ corresponds with the typical deviations of the measurements from the mean value becoming smaller, this confirms the earlier observation that the mean value of a set of measurements gets closer to the true value as σ decreases.

If the standard deviation is used as a unit of error, the Gaussian curve can be used to determine the probability that the deviation in any particular measurement in a Gaussian data set is greater than a certain value. By substituting the expression for $F(D)$ in (3.12) into the probability equation (3.9), the probability that the error lies in a band between error levels D_1 and D_2 can be expressed as:

$$P(D_1 \leq D \leq D_2) = \int_{D_1}^{D_2} \frac{1}{\sigma\sqrt{2\pi}} e^{(-D^2/2\sigma^2)}\, \mathrm{d}D \tag{3.13}$$

Solution of this expression is simplified by the substitution:

$$z = D/\sigma \tag{3.14}$$

The effect of this is to change the error distribution curve into a new Gaussian distribution that has a standard deviation of one ($\sigma = 1$) and a mean of zero. This new form, shown in Figure 3.7, is known as a *standard Gaussian curve*, and the dependent

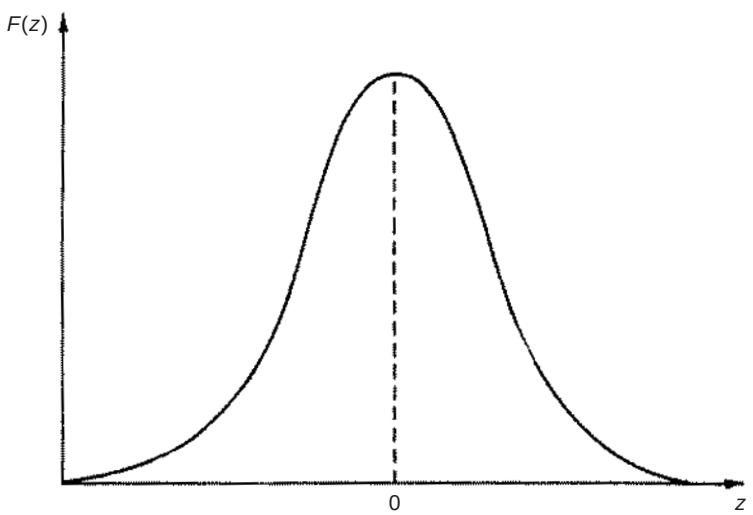

Fig. 3.7 Standard Gaussian curve ($F(z)$ versus z).

variable is now z instead of D. Equation (3.13) can now be re-expressed as:

$$P(D_1 \leq D \leq D_2) = P(z_1 \leq z \leq z_2) = \int_{z_1}^{z_2} \frac{1}{\sigma\sqrt{2\pi}} e^{(-z^2/2)}\,\mathrm{d}z \tag{3.15}$$

Unfortunately, neither equation (3.13) nor (3.15) can be solved analytically using tables of standard integrals, and numerical integration provides the only method of solution. However, in practice, the tedium of numerical integration can be avoided when analysing data because the standard form of equation (3.15), and its independence from the particular values of the mean and standard deviation of the data, means that standard Gaussian tables that tabulate $F(z)$ for various values of z can be used.

Standard Gaussian tables

A standard Gaussian table, such as that shown in Table 3.1, tabulates $F(z)$ for various values of z, where $F(z)$ is given by:

$$F(z) = \int_{-\infty}^{z} \frac{1}{\sigma\sqrt{2\pi}} e^{(-z^2/2)}\,\mathrm{d}z \tag{3.16}$$

Thus, $F(z)$ gives the proportion of data values that are less than or equal to z. This proportion is the area under the curve of $F(z)$ against z that is to the left of z. Therefore, the expression given in (3.15) has to be evaluated as $[F(z_2) - F(z_1)]$. Study of Table 3.1 shows that $F(z) = 0.5$ for $z = 0$. This confirms that, as expected, the number of data values ≤ 0 is 50% of the total. This must be so if the data only has random errors. It will also be observed that Table 3.1, in common with most published standard Gaussian tables, only gives $F(z)$ for positive values of z. For negative values of z, we can make use of the following relationship because the frequency distribution curve is normalized:

$$F(-z) = 1 - F(z) \tag{3.17}$$

($F(-z)$ is the area under the curve to the left of $(-z)$, i.e. it represents the proportion of data values $\leq -z$.)

Example 3.3

How many measurements in a data set subject to random errors lie outside deviation boundaries of $+\sigma$ and $-\sigma$, i.e. how many measurements have a deviation greater than $|\sigma|$?

Solution

The required number is represented by the sum of the two shaded areas in Figure 3.8. This can be expressed mathematically as:

$$P(E < -\sigma \text{ or } E > +\sigma) = P(E < -\sigma) + P(E > +\sigma)$$

For $E = -\sigma$, $z = -1.0$ (from equation 3.12).

Using Table 3.1:

$$P(E < -\sigma) = F(-1) = 1 - F(1) = 1 - 0.8413 = 0.1587$$

Table 3.1 Standard Gaussian table

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
					$F(z)$					
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7793	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8906	0.8925	0.8943	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9648	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9926	0.9928	0.9930	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9986	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

Similarly, for $E = +\sigma, z = +1.0$, Table 3.1 gives:

$$P(E > +\sigma) = 1 - P(E < +\sigma) = 1 - F(1) = 1 - 0.8413 = 0.1587.$$

(This last step is valid because the frequency distribution curve is normalized such that the total area under it is unity.)

Thus

$$P[E < -\sigma] + P[E > +\sigma] = 0.1587 + 0.1587 = 0.3174 \sim 32\%$$

i.e. 32% of the measurements lie outside the $\pm\sigma$ boundaries, then 68% of the measurements lie inside.

The above analysis shows that, for Gaussian-distributed data values, 68% of the measurements have deviations that lie within the bounds of $\pm\sigma$. Similar analysis shows

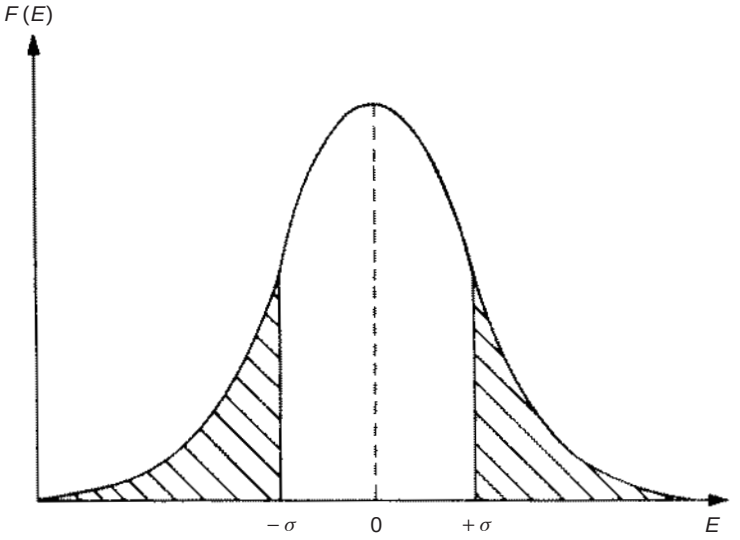

Fig. 3.8 $\pm\sigma$ boundaries.

that boundaries of $\pm 2\sigma$ contain 95.4% of data points, and extending the boundaries to $\pm 3\sigma$ encompasses 99.7% of data points. The probability of any data point lying outside particular deviation boundaries can therefore be expressed by the following table.

Deviation boundaries	*% of data points within boundary*	*Probability of any particular data point being outside boundary*
$\pm\sigma$	68.0	32.0%
$\pm 2\sigma$	95.4	4.6%
$\pm 3\sigma$	99.7	0.3%

Standard error of the mean

The foregoing analysis has examined the way in which measurements with random errors are distributed about the mean value. However, we have already observed that some error remains between the mean value of a set of measurements and the true value, i.e. averaging a number of measurements will only yield the true value if the number of measurements is infinite. If several subsets are taken from an infinite data population, then, by the central limit theorem, the means of the subsets will be distributed about the mean of the infinite data set. The error between the mean of a finite data set and the true measurement value (mean of the infinite data set) is defined as the *standard error of the mean*, α. This is calculated as:

$$\alpha = \sigma/\sqrt{n} \tag{3.18}$$

α tends towards zero as the number of measurements in the data set expands towards infinity. The measurement value obtained from a set of n measurements, $x_1, x_2, \cdots x_n$,

can then be expressed as:

$$x = x_{\text{mean}} \pm \alpha$$

For the data set C of length measurements used earlier, $n = 23$, $\sigma = 1.88$ and $\alpha = 0.39$. The length can therefore be expressed as 406.5 ± 0.4 (68% confidence limit). However, it is more usual to express measurements with 95% confidence limits ($\pm 2\sigma$ boundaries). In this case, $2\sigma = 3.76$, $2\alpha = 0.78$ and the length can be expressed as 406.5 ± 0.8 (95% confidence limits).

Estimation of random error in a single measurement

In many situations where measurements are subject to random errors, it is not practical to take repeated measurements and find the average value. Also, the averaging process becomes invalid if the measured quantity does not remain at a constant value, as is usually the case when process variables are being measured. Thus, if only one measurement can be made, some means of estimating the likely magnitude of error in it is required. The normal approach to this is to calculate the error within 95% confidence limits, i.e. to calculate the value of the deviation D such that 95% of the area under the probability curve lies within limits of $\pm D$. These limits correspond to a deviation of $\pm 1.96\sigma$. Thus, it is necessary to maintain the measured quantity at a constant value whilst a number of measurements are taken in order to create a reference measurement set from which σ can be calculated. Subsequently, the maximum likely deviation in a single measurement can be expressed as: Deviation $= \pm 1.96\sigma$. However, this only expresses the maximum likely deviation of the measurement from the calculated mean of the reference measurement set, which is not the true value as observed earlier. Thus the calculated value for the standard error of the mean has to be added to the likely maximum deviation value. Thus, the maximum likely error in a single measurement can be expressed as:

$$\text{Error} = \pm(1.96\sigma + \alpha) \tag{3.19}$$

Example 3.4

Suppose that a standard mass is measured 30 times with the same instrument to create a reference data set, and the calculated values of σ and α are $\sigma = 0.43$ and $\alpha = 0.08$. If the instrument is then used to measure an unknown mass and the reading is 105.6 kg, how should the mass value be expressed?

Solution

Using (3.19), $1.96\sigma + \alpha = 0.92$. The mass value should therefore be expressed as: 105.6 ± 0.9 kg.

Before leaving this matter, it must be emphasized that the maximum error specified for a measurement is only specified for the confidence limits defined. Thus, if the maximum error is specified as $\pm 1\%$ with 95% confidence limits, this means that there is still 1 chance in 20 that the error will exceed $\pm 1\%$.

Distribution of manufacturing tolerances

Many aspects of manufacturing processes are subject to random variations caused by factors that are similar to those that cause random errors in measurements. In most cases, these random variations in manufacturing, which are known as *tolerances*, fit a

3.7 Self-test questions

3.1 Explain the difference between systematic and random errors. What are the typical sources of these two types of error?

3.2 In what ways can the act of measurement cause a disturbance in the system being measured?

3.3 Suppose that the components in the circuit shown in Figure 3.1(a) have the following values:

$$R_1 = 330\,\Omega; \quad R_2 = 1000\,\Omega; \quad R_3 = 1200\,\Omega; \quad R_4 = 220\,\Omega; \quad R_5 = 270\,\Omega.$$

If the instrument measuring the output voltage across AB has a resistance of 5000 Ω, what is the measurement error caused by the loading effect of this instrument?

3.4 Instruments are normally calibrated and their characteristics defined for particular standard ambient conditions. What procedures are normally taken to avoid measurement errors when using instruments that are subjected to changing ambient conditions?

3.5 The voltage across a resistance R_5 in the circuit of Figure 3.10 is to be measured by a voltmeter connected across it.

(a) If the voltmeter has an internal resistance (R_m) of 4750 Ω, what is the measurement error?

(b) What value would the voltmeter internal resistance need to be in order to reduce the measurement error to 1%?

3.6 In the circuit shown in Figure 3.11, the current flowing between A and B is measured by an ammeter whose internal resistance is 100 Ω. What is the measurement error caused by the resistance of the measuring instrument?

3.7 What steps can be taken to reduce the effect of environmental inputs in measurement systems?

3.8 The output of a potentiometer is measured by a voltmeter having a resistance R_m, as shown in Figure 3.12. R_t is the resistance of the total length X_t of the potentiometer and R_i is the resistance between the wiper and common point C for a general wiper position X_i. Show that the measurement error due to the resistance

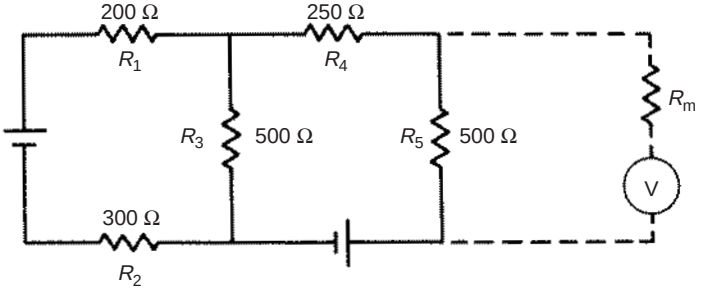

Fig. 3.10 Circuit for question 3.5.

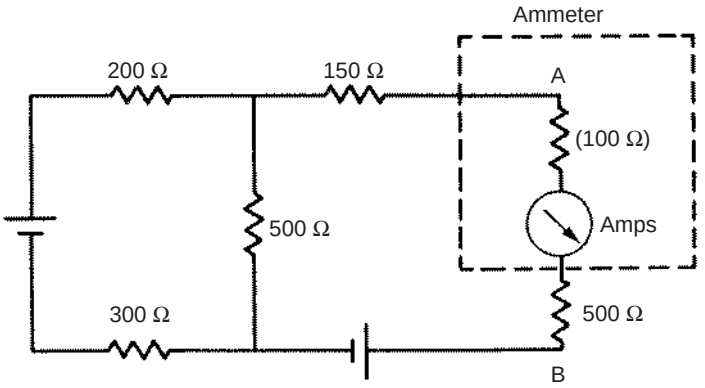

Fig. 3.11 Circuit for question 3.6.

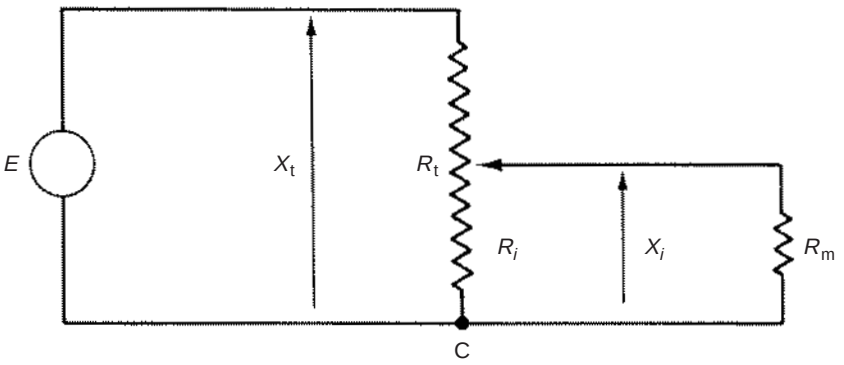

Fig. 3.12 Circuit for question 3.8.

R_m of the measuring instrument is given by:

$$\text{Error} = E\frac{R_i^2(R_t - R_i)}{R_t(R_iR_t + R_mR_t - R_i^2)}$$

Hence show that the maximum error occurs when X_i is approximately equal to $2X_t/3$. (Hint – differentiate the error expression with respect to R_i and set to 0. Note that maximum error does not occur exactly at $X_i = 2X_t/3$, but this value is very close to the position where the maximum error occurs.)

3.9 In a survey of 15 owners of a certain model of car, the following figures for average petrol consumption were reported.

25.5 30.3 31.1 29.6 32.4 39.4 28.9 30.0 33.3 31.4 29.5 30.5 31.7 33.0 29.2

Calculate the mean value, the median value and the standard deviation of the data set.

3.10 (a) What do you understand by the term *probability density function*?
(b) Write down an expression for a Gaussian probability density function of given mean value μ and standard deviation σ and show how you would obtain the best estimate of these two quantities from a sample of population n.
(c) The following ten measurements are made of the output voltage from a high-gain amplifier that is contaminated due to noise fluctuations:

$$1.53,\ 1.57,\ 1.54,\ 1.54,\ 1.50,\ 1.51,\ 1.55,\ 1.54,\ 1.56,\ 1.53$$

Determine the mean value and standard deviation. Hence estimate the accuracy to which the mean value is determined from these ten measurements. If one thousand measurements were taken, instead of ten, but σ remained the same, by how much would the accuracy of the calculated mean value be improved?

3.11 The following measurements were taken with an analogue meter of the current flowing in a circuit (the circuit was in steady state and therefore, although the measurements varied due to random errors, the current flowing was actually constant):

$$21.5\,\text{mA},\ 22.1\,\text{mA},\ 21.3\,\text{mA},\ 21.7\,\text{mA},\ 22.0\,\text{mA},\ 22.2\,\text{mA},\ 21.8\,\text{mA},$$
$$21.4\,\text{mA},\ 21.9\,\text{mA},\ 22.1\,\text{mA}$$

Calculate the mean value, the deviations from the mean and the standard deviation.

3.12 The measurements in a data set are subject to random errors but it is known that the data set fits a Gaussian distribution. Use standard Gaussian tables to determine the percentage of measurements that lie within the boundaries of $\pm 1.5\sigma$, where σ is the standard deviation of the measurements.

3.13 The thickness of a set of gaskets varies because of random manufacturing disturbances but the thickness values measured belong to a Gaussian distribution. If the mean thickness is 3 mm and the standard deviation is 0.25, calculate the percentage of gaskets that have a thickness greater than 2.5 mm.

3.14 A 3 volt d.c. power source required for a circuit is obtained by connecting together two 1.5 V batteries in series. If the error in the voltage output of each battery is specified as ±1%, calculate the likely maximum possible error in the 3 volt power source that they make up.

3.15 In order to calculate the heat loss through the wall of a building, it is necessary to know the temperature difference between the inside and outside walls. If temperatures of 5°C and 20°C are measured on each side of the wall by mercury-in-glass thermometers with a range of 0°C to +50°C and a quoted inaccuracy figure of ±1% of full-scale reading, calculate the likely maximum possible error in the calculated figure for the temperature difference.

3.16 The power dissipated in a car headlight is calculated by measuring the d.c. voltage drop across it and the current flowing through it ($P = V \times I$). If the possible errors in the measured voltage and current values are ±1% and ±2% respectively, calculate the likely maximum possible error in the power value deduced.

3.17 The resistance of a carbon resistor is measured by applying a d.c. voltage across it and measuring the current flowing ($R = V/I$). If the voltage and current values

are measured as 10 ± 0.1 V and 214 ± 5 mA respectively, express the value of the carbon resistor.

3.18 The density (d) of a liquid is calculated by measuring its depth (c) in a calibrated rectangular tank and then emptying it into a mass measuring system. The length and width of the tank are (a) and (b) respectively and thus the density is given by:

$$d = m/(a \times b \times c)$$

where m is the measured mass of the liquid emptied out.
If the possible errors in the measurements of a, b, c and m are 1%, 1%, 2% and 0.5% respectively, determine the likely maximum possible error in the calculated value of the density (d).

3.19 The volume flow rate of a liquid is calculated by allowing the liquid to flow into a cylindrical tank (stood on its flat end) and measuring the height of the liquid surface before and after the liquid has flowed for 10 minutes. The volume collected after 10 minutes is given by:

$$\text{Volume} = (h_2 - h_1)\pi(d/2)^2$$

where h_1 and h_2 are the starting and finishing surface heights and d is the measured diameter of the tank.

(a) If $h_1 = 2$ m, $h_2 = 3$ m and $d = 2$ m, calculate the volume flow rate in m^3/min.
(b) If the possible error in each measurement h_1, h_2 and d is $\pm 1\%$, determine the likely maximum possible error in the calculated value of volume flow rate.

References and further reading

ANSI/ASME standards (1985) ASME performance test codes, supplement on instruments and apparatus, part 1: measurement uncertainty, American Society of Mechanical Engineers, New York.

Bennington, P.R. and Robinson, D.K. (1992) *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill.

Caulcott, E. (1973) *Significance Tests*, Routledge and Kegan Paul.

Miller, I.R., Freung, J.E. and Johnson, R. (1990) *Probability and Statistics for Engineers*, Prentice-Hall.

Topping, J. (1962) *Errors of Observation and Their Treatment*, Chapman and Hall.