

Engineering Experimentation

Lecture Notes

1.Introduction	1
1.1. APPLICATIONS OF ENGINEERING EXPERIMENTATION AND MEASUREMENT	1
1.1.1. Measurement in Engineering Experimentation	2
1.1.2. Measurement in Operational Systems	3
1.2. OBJECTIVE AND OVERVIEW	4
2. General Characteristics of Measurement Systems	5
2.1. GENERALIZED MEASUREMENT SYSTEM	5
2.2. VALIDITY OF MEASUREMENT.....	6
2.2.1. Measurement Error and Related Definitions.....	7
2.2.2. Calibration of Measurement Systems.....	14
2.3. DYNAMIC MEASUREMENTS.....	22

1.Introduction

Experimentation is the backbone of modern physical science. In engineering, carefully designed experiments are needed to conceive and verify theoretical concepts, develop new methods and products, commission sophisticated new engineering systems, and evaluate the performance and behavior of existing products. Experimentation and the design of measurement systems are major engineering activities. In this chapter we give an overview of the applications of experiments and measurement systems and describe briefly how this book will prepare the reader for professional activities in these areas.

1.1. APPLICATIONS OF ENGINEERING EXPERIMENTATION AND MEASUREMENT

Engineering measurement applications can broadly be broken into two categories. The first of these is measurement in engineering experimentation, in which new information is being sought, and the second is measurement in operational devices for monitoring and control purposes.

1.1.1. Measurement in Engineering Experimentation

Engineering experimentation, which in a general sense involves using the measurement process to seek new information, ranges in scope from experiments to establish new concepts all the way to testing of existing products to determine maintenance requirements. Such experimentation falls broadly into three categories:

1. Research experimentation
2. Development experimentation
3. Performance testing

The primary difference between research and development is that in the former, concepts for new products or processes are being sought (often unsuccessfully), while in the latter, known concepts are being used to establish potential commercial products.

Carbon-fiber composites represent a relatively recent example of the research and development process. Carbon-fiber composites are now used commercially for such diverse products as golf clubs and aircraft control surfaces. In the research phase, methods were suggested and evaluated to produce carbon fibers in small quantities and tests were performed to determine the physical properties of samples. The results of the research activities were so promising that many development activities were initiated. These activities included development of large-scale fiber manufacturing processes and development of methods to fabricate fiber composite parts. Although there are now many products using carbon fibers, developmental activities in this area continue. The fuselage of the commercial airliner, Boeing 787, is constructed entirely from carbon fiber material and this advance saved considerable weight, resulting in improved efficiency, and is considered a major advance in aircraft technology.

Research experiments are frequently highly uncertain and often lead to dead ends. The risk is high, either because the experiment itself may be unsuccessful or because the experimental result may not be as wanted. Research experimentation is usually performed in universities or special research organizations. On the other hand, development programs usually have better defined goals than research programs and frequently result in an operational product. Sometimes, however, the product will not be deemed competitive and will never be produced in quantity. Development programs are usually performed by product manufacturers.

Although the instrumentation must function properly during the research or development program, it may be delicate and require considerable attention. Special measurement techniques may be created. Experimental measuring systems whose characteristics are not completely defined may also be suitable for such testing programs. The engineers and scientists performing such tests are generally sophisticated in the fine points of the instruments and can compensate for deficiencies.

Performance testing is somewhat different from research and development experimental activities. Performance testing is done on products that have been developed and in many cases are already on the market. Performance testing may be carried out to demonstrate

applicability for a particular application, to assess reliability, or to determine product lifetime. This testing may be done either by the manufacturer, the supplier, the customer, or an independent laboratory. As an example, a performance test might be used to demonstrate that an electronic device which functions satisfactorily in a stationary environment will also function in an aircraft application with high levels of vibration.

Another type of performance testing is the periodic testing of operating devices to determine needs for maintenance. Utilities normally perform this type of testing in power plants to make sure that the efficiencies of large pumps, heat exchangers, and other components are adequate. Poor performance will lead to maintenance actions. Instruments may be in place for such tests, but they may need repair, and supplementary instruments may be required at the time of the tests. Commissioning of process plants may also involve extensive but standardized testing to demonstrate conformance to design specifications.

Measuring systems for performance testing are generally established devices with well-defined characteristics. The systems need to be reliable, and significant interpretation of ambiguities in the measured values should not be required since the people performing the tests are often technicians.

Often, professional engineering organizations such as the American Society of Mechanical Engineers (ASME), the Institute of Electrical and Electronic Engineers (IEEE), and the International Society of Automation (ISA) have established detailed procedures for performance testing.

1.1.2. Measurement in Operational Systems

Many dynamic systems are instrumented for monitoring or control purposes. Such systems range from simple furnaces for home heating to extremely complex jet aircraft. One very sophisticated but everyday measurement and control system is the engine control system of modern automobiles. These systems have sensors to measure variables such as airflow, engine speed, water temperature, and exhaust gas composition and use a computer to determine the correct fuel flow rate. These engine control systems are very compact and are specially engineered for the particular application.

Elaborate measurement and control systems are needed in complex process plants such as oil refineries, steam power plants, and sewage treatment facilities. Such systems may have hundreds of sensors and use computers to collect and interpret the data and control the process. This particular class of applications is so large that it is a specialized field in its own right, called process control. While the complete measuring systems for such applications are specifically engineered, the components are generally modular and standardized.

Instrumentation for operating systems must be very durable and reliable. Sensors that need to be calibrated very frequently would present major problems in these applications. In many cases, the measuring systems have to be designed such that by redundancy or other techniques, a failed component can be readily identified so that the operating system can continue to operate correctly or at least be safely shut down.

1.2. OBJECTIVE AND OVERVIEW

The objective of this book is to provide the reader with the skills necessary to perform an engineering experiment systematically—from the definition of the experimental need to the completion of the final report. A systematic approach includes careful planning and analytical design of the experiment before it is constructed, demonstration of the validity of the test apparatus, analysis of the test results, and reporting of the final results. The emphasis is on the design of the experiment and the analysis of the results; however, guidance is given on other activities. Chapters 2 through 11 provide the technical information necessary to design an experimental system and interpret the results. This information is also applicable to the design of the measurement (but not control) systems of process plants. Chapter 12 provides an overview of the overall experimental design process and provides guidelines on planning, designing, scheduling, and documenting experimental projects.

1.3. DIMENSIONS AND UNITS

The International System of Units (SI) is the most widely used unit system in the world, due to its consistency and simplicity. However, in the United States and some other countries, a unit system based on the old British unit system is still widely used. Product specifications and data tables are frequently given in British units. For example, the range of a pressure measurement device might be specified in pounds per square inch (psi). To assist the reader in developing capabilities in both unit systems, both SI and British units systems are used in example problems in this book.

The physical world is described with a set of dimensions. Length, mass, time, and temperature are dimensions. When a numerical value is assigned to a dimension, it must be done in a unit system. For example, we can describe the temperature (dimension) of an ice-water mixture in either the SI unit system (0°C) or the British unit system (32°F). International conferences have established a set of SI base units. Table 1.1 lists the base SI units and the corresponding British units. There are two additional base units, the candela for light intensity and the mole for the amount of a substance, but these units are not used in this book. Each of these base units has a corresponding standard such that all users of the unit can compare their results. Standards are discussed in Chapter 2.

Other engineering quantities, such as force and power, are related to the dimensions of the base units through physical laws and definitions. The dimension of force is defined by Newton's second law: $F = m a / g_c$

where F is force, m is mass, and a is acceleration. g_c is a proportionality constant, which depends on the unit system. In the SI unit system, the unit of force is the newton (N) and is defined as the force exerted on a 1 kg mass to generate an acceleration of 1 m/s^2 . In the SI system, g_c has a value of unity. In the British unit system, the unit of force is the pound force (lbf) and is defined as the force exerted on a 1 lb mass at the standard gravitational acceleration of 32.174 ft/sec^2 . In this case, the value of g_c has to be taken as $32.174 \text{ lbf-ft/lbf-sec}^2$.

In this book, in equations based on Newton's second law, the constant g_c is taken to be unity and does not appear in the equations. All equations will produce a dimensionally correct result if the SI system of units is used properly. Sometimes in the British system, mass is specified using a unit called the slug, defined as 32.174 lbf . When the slug is used to define mass, the

constant g_c is also unity. Unfortunately, the slug is not a widely used unit, and most British-unit data tables and specifications use lbm for the mass unit. Consequently, mass numbers supplied in lbm must be converted by dividing by the constant 32.174 when using them in equations in this book. Another characteristic of the British unit system is that two units are used for energy. In mechanical systems, the unit of energy is the ft-lbf, while in thermal systems, it is the Btu. The conversion factor is $1 \text{ Btu} = 778 \text{ ft-lbf}$.

TABLE 1.1 Base SI and British Units		
Dimension	SI unit	British unit
Mass	kilogram (kg)	pound mass (lbm)
Length	meter (m)	foot (ft)
Time	second (s)	second (s)
Temperature	Kelvin (K)	Rankine degree ($^{\circ}\text{R}$)
Electric current	ampere (A)	ampere (A)

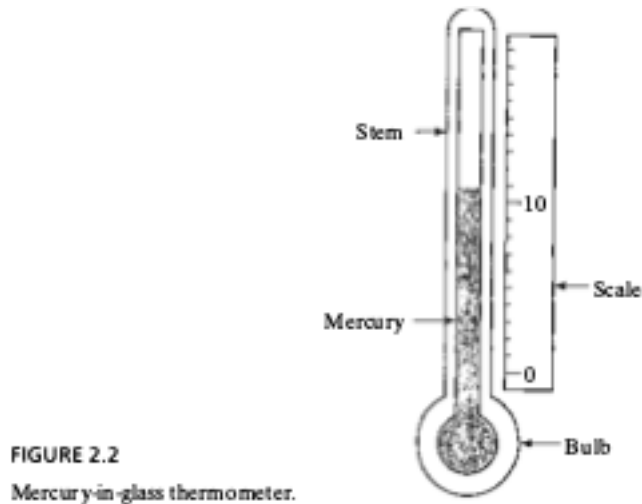
When using any unit system, great care is required to make sure that the units are consistent, particularly with British units. It is recommended that a units check be performed for all calculations using British units.

2. General Characteristics of Measurement Systems

A necessary part of planning an experiment is to determine the specifications for the required measurement systems. The characteristics of many specific measuring devices are detailed in Chapters 8 to 10. In this chapter, significant general characteristics of measurement systems are described and definitions are provided for common descriptive terms.

2.1. GENERALIZED MEASUREMENT SYSTEM

In any experiment the experimenter seeks to obtain numerical values for certain physical variables. These unknown variables are known as **measurands**. Examples of measurands include temperature, velocity, and voltage. The measurement system senses the measurand and produces a unique numerical value that describes the measurand. In most cases, the measurement system can be viewed as consisting of three subsystems. As shown in Figure 2.1, these three subsystems are the sensing element, the signal modification subsystem, and the recording or indicating device. The sensing element has a significant physical characteristic that changes in response to changes in the measurand. The signal modification subsystem changes the output of the sensing element in some way to make it more suitable for the indicating or recording device. If the user simply reads the output and perhaps copies it to paper, the final device is an indicator. If the output value is saved automatically in some way, the final device is a recorder. A measurement system may have both an indicating device and a recording device. In modern measurement systems, this final stage is often a computer, which can not only display and record the data but can also manipulate the data mathematically.



A simple example of a measurement system is a common mercury-in-glass thermometer (Figure 2.2), which could be used to measure the temperature of water in a container. In this device, the volume of the mercury in the bulb depends on the temperature of the mercury. If the bulb has been in contact with the water for a sufficient time, the mercury will have the same temperature as the surrounding fluid. Hence a measurement of the volume of the mercury can be used to determine the temperature of the water. Unfortunately, it is very difficult to measure the small change in volume of the mercury. If the mercury had the shape of a sphere, the change in diameter would be very small. Therefore, signal modification is required. For the thermometer, signal modification is accomplished by connecting the bulb to the stem. The inside diameter of the stem is very small relative to the diameter of the bulb, and although the change in mercury volume is small, this small change in volume produces a large change in length of the stem mercury column. Actually, the displacement of mercury in the stem is proportional to the differential thermal expansion between the mercury and the glass envelope. Finally, an indicating device is required. In the case of the thermometer, this is accomplished with a scale that is either next to the glass stem or engraved on it directly.

These three subsystems are quite obvious in most measuring devices. This is particularly true for modern measurement systems using electrical output-sensing devices, in which the three subsystems are often physically separate devices. There are, however, some common measuring systems in which all three subsystems are difficult to identify or the components are combined.

2.2. VALIDITY OF MEASUREMENT

It is very important to the experimenter that the output of a measurement system truly states the actual value of the measurand. That is, the experimenter must be convinced that the output of the measurement system is valid. Of course, no measurement system is perfect—there will always be some deviation between the actual value of the measurand and the measurement system output. This deviation must simply be small enough that the output can

be used for its intended purpose. Generally speaking, the smaller the allowed deviation, the more expensive will be the measurement system.

2.2.1. Measurement Error and Related Definitions

Several standard terms are used to specify the validity of a measurement. The error of a measurement is defined as the difference between the measured value and the true value of the measurand:

$$\text{Error} = \text{measured value} - \text{true value}$$

Error in this technical usage does not imply that there is any mistake in the measurement process, although mistakes can cause errors. Normally, the experimenter can never really know the error of a measurement. If the true value of the measurand were known, there would be no need to make the measurement (except in the process of calibration, where measurements are made of measurands whose values are independently known). What the experimenter can estimate, however, is the uncertainty interval (or simply uncertainty) of the measurement. The uncertainty is an estimate (with some level of confidence) of the limits of error in the measurement. For example, it might be stated that with 95 % confidence, the uncertainty of a voltage measurement is ± 1 volt. This means that the error will be greater than 1 V in less than 5 % of the cases where we have made such uncertainty predictions. Narrow uncertainty intervals are usually achieved by using calibrated, high-quality measuring systems.

Errors in experiments generally fall into two categories: systematic errors (sometimes called fixed or bias errors) and random errors (sometimes called precision errors). Although both types of error degrade the validity of the data, their causes are different and different actions must be taken to minimize them.

Systematic errors are consistent, repeatable errors. For example, a measuring system might give a consistent 10% high reading. In other cases, the output might be the same absolute amount low for all readings. In general, if the same measuring system is used in the same way more than once to measure the same value of the measurand, the systematic error will be the same each time.

The first major source of systematic error is that resulting from calibration of the measurement system. If the calibration process has some error, that error will be carried into the measurement as a systematic error. Even the most exact calibration will result in a residual systematic error. These are known as calibration errors. One very common source of calibration systematic error is nonlinearity. Many modern systems are treated as if they have a linear relationship between the input and the output, and the actual nonlinearity of the system will cause errors.

The second major source of systematic error results from the use of a measuring system in a particular application where the insertion of the measuring device alters the measurand. For example, connecting a temperature-measuring device to a surface may in fact change the local temperature of the surface. Such errors are known as loading errors. As another example, consider placing a mercury-in-glass thermometer into a beaker of water. If the beaker and the thermometer are initially at different temperatures, energy will be exchanged between them, and the measured temperature will be neither the initial water temperature nor the initial thermometer temperature (but usually closer to the water temperature). The thermometer is an intrusive measurement device and produces a significant loading error. Some measuring

devices with negligible loading errors are called nonintrusive. For example, devices are available that measure temperature by sensing the infrared radiation emitted. Such a device would have a negligible effect on the measured temperature and is said to be nonintrusive.

A third major systematic error results because the measuring system is affected by variables other than the measurand. For example, a thermometer used to measure the air temperature in a room will read too low, due to thermal radiation effects, if the walls are cooler than the air. A related error is the spatial error. If the measurand varies in a spatial region and yet a single measurement or a limited number of measurements are used to determine the average value for the region, there will be a spatial error.

Systematic errors are often not obvious to the experimenter-the measuring system will show clear and consistent changes in output following changes in the measurand, yet it will still have significant error. In setting up an experiment, considerable time may be required to detect and minimize systematic errors. Systematic errors in the measuring system may be detected and reduced by the process of calibration, discussed later. Some systematic errors caused by using the measuring system in a particular application may be reduced by analytical correction of the data for unwanted effects.

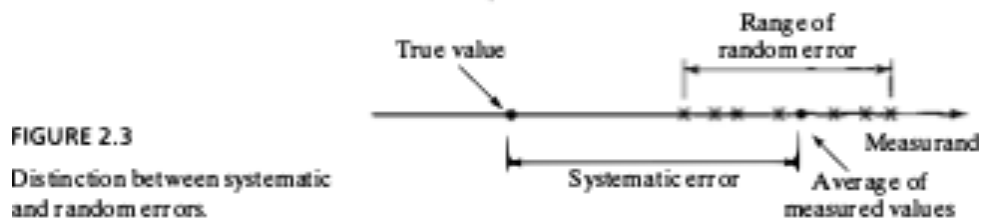
Random errors are those caused by a lack of repeatability in the output of the measuring system. The distinction between systematic and random errors is shown graphically in Figure 2.3. The scatter in the data represents random error, and the deviation between the average of the readings and true value demonstrates the systematic error. The random error in a single measurement can be estimated as the difference between the single reading and the average of all readings of the same value of measurand:

$$\text{Random error} = \text{reading} - \text{average of readings}$$

which distinguishes the random error from the systematic error. The systematic error can be estimated by using the following equation:

$$\text{systematic error} = \text{average of readings} - \text{true value}$$

For these estimates of systematic and random errors to be reasonable, the number of readings forming the average must be large enough to eliminate the effects of random error in individual measurements on the average.



Random errors can originate from the measuring system itself, from the experimental system, or from the environment. Random errors are usually caused by uncontrolled variables in the measurement process. For example, the performance of an amplifier may be slightly sensitive to its temperature. If we do not measure or control the temperature, performance measurements may show a certain variability or scatter. One very important environmental cause of random error is electrical noise. Experiments and measuring systems normally

operate in a sea of electric and magnetic fields caused by sources such as building wiring and local radio stations. These electric and magnetic background fields can affect readings by randomly altering voltages in measuring systems and connecting wiring.

Random errors can often be minimized by eliminating uncontrolled variables or properly shielding or grounding the measuring system. Remaining random errors may be amenable to statistical analysis—for example, a large number of readings can be averaged. Statistical analysis of data is discussed in detail in Chapter 6. Example 2.1 shows how to estimate the systematic and maximum random errors for a calibration test of a voltmeter.

Example 2.1.

In a calibration test, 10 measurements using a digital voltmeter have been made of the voltage of a battery that is known to have a true voltage of 6.11 V. The readings are: 5.98, 6.05, 6.10, 6.06, 5.99, 5.96, 6.02, 6.09, 6.03, and 5.99 V. Estimate the systematic and maximum random errors caused by the voltmeter.

Solution: First, determine the average of the 10 readings:

$$\text{average } V = 6.03$$

Then the estimate of the systematic error is computed as follows:

$$\text{systematic error} = \text{average value} - \text{true value} = 6.03 - 6.11 = -0.08 \text{ V}$$

To estimate the maximum random error, we need to determine the reading that deviates the most from the average reading. This is the reading of 5.96 V. The maximum precision error is thus

$$\text{maximum random error} = 5.96 - 6.03 = -0.07 \text{ V}$$

Comment: It should be noted that this simple statement of maximum random error may not adequately describe random errors in a measuring system. For example, it may be based on a single bad reading. Statistical methods described in Chapters 6 and 7 provide procedures to determine random errors, which include all of the readings and also provide a basis to eliminate certain bad data.

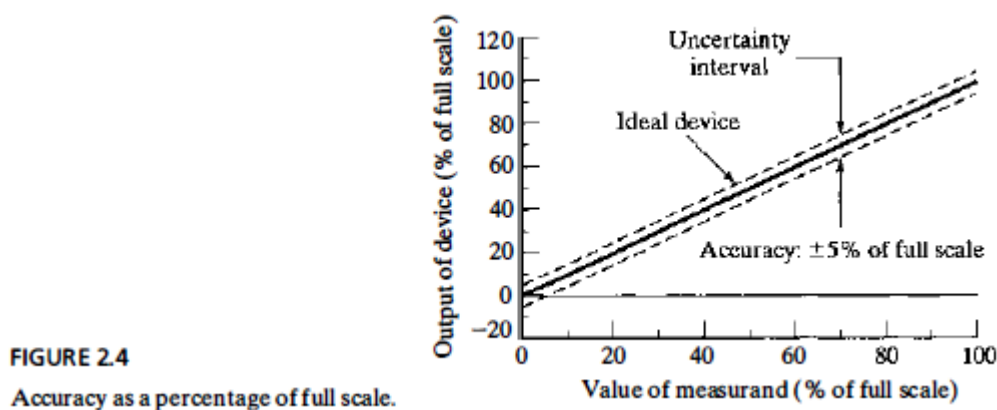
A measuring system is only designed to operate over a specified **range** of measurands. The range of a measuring system describes the values of the measurand to which that measuring system will respond properly—values of the measurand outside the range are not expected to produce useful output. For example, a voltmeter may have a range of 0 to 10 V and would not give a correct response to measurands of -5 or 13 V. The span of a measuring system is the difference between the upper and lower values of the range. For example, a voltmeter with a range of $\pm 3\text{V}$ has a span of 6 V.

Accuracy, defined as the closeness of agreement between a measured value and the true value, is a common term used to specify uncertainty. Measuring device manufacturers frequently state a value for accuracy as part of the device specifications. Although the term accuracy is generally used, it is really the inaccuracy that is specified. As commonly used, manufacturer specifications of accuracy describe residual uncertainty that exists when a

device has been properly adjusted and calibrated and is used in a specified manner. Accuracy specifications generally include residual systematic and random errors in the measuring system itself. Accuracy might be given for either a component of a measuring system (e.g., a sensor) or for a complete system and is most often specified as a percentage of full-scale output. For example, if the output of a device can range from 0 to 5 V, and the accuracy is stated as $\pm 5\%$ of full scale, the uncertainty is ± 0.25 V, regardless of the reading. The procedure to determine accuracy is described in Section 2.2.2 and is based on information in ANSI/ISA (1979). If more than one component is used in the measurement of a single measurand, a combined uncertainty must be determined. Methods to estimate overall or total uncertainty are described in Chapter 7.

As Figure 2.4 shows for a typical measuring device with an accuracy of $\pm 5\%$ of full scale, at readings below full scale, the percent uncertainty in the reading will be greater than 5%. At readings toward the lower end of the range, the percent uncertainty might be completely unsatisfactory. This problem with high uncertainty at the low end of the range is a major concern in selecting a measuring system. To minimize uncertainty, the experimenter should select measuring systems such that important readings will fall in the middle to upper portions of the range. For example, it would adversely affect uncertainty if a 0-to- 200°C thermometer was used to measure a room temperature around 20°C . A 0 to 30°C thermometer would be far more appropriate.

There are other statements of accuracy, such as an accuracy stated as a percent of reading. Manufacturers of high-quality measuring systems and components will normally give enough information about their products so that the experimenter can determine the uncertainty in the measurement that is due to the measuring system itself. The experimenter may have to enlarge the uncertainty interval to account for other error sources that result from the specific application.



Precision is another term frequently used to describe a measuring system or component and characterizes the random error. A highly precise measuring system will give the same value each time it is read, but it may not be very accurate—it may simply give the same inaccurate answer each time the measurement is made. In general, the accuracy of a measuring system cannot be any better than the measurement constraints provided by the instrument precision (although it can be much worse). Accuracy and precision are overall characteristics that describe the validity of measurement. Each characteristic is determined by a number of specific sources of uncertainty (errors).

In measuring devices, accuracy is often degraded by a phenomenon known as **hysteresis**. As shown in Figure 2.5, for the same value of the measurand, different out- put readings may be

obtained if the measurand was increasing prior to taking the reading than if the measurand was decreasing. Hysteresis is caused by such effects as friction, mechanical flexure of internal parts, and electrical capacitance. Errors due to hysteresis are known as **hysteresis errors**. If a measurement is repeated in exactly the same manner, errors due to hysteresis would be repeatable and hence would be considered systematic errors. However, in common measuring processes, the experimenter generally may not know if the measurand was increasing or decreasing when a measurement was made. Hence, the effect of hysteresis will appear random. However, when estimating total uncertainty, it is normally conservative to treat hysteresis errors as a systematic error. Hysteresis error is usually a component of the instrument manufacturers' specification of accuracy.

Another important characteristic of a measuring system is the **resolution**. If a measurand is varied continuously, many measurement devices will show an output that changes in discrete steps. This inability of the measurement system to follow changes in the measurand exactly results in a **resolution error**. Resolution is usually treated as a random error. Internal characteristics of a measuring system may limit resolution. The sensing element itself may not produce a continuous output with a smoothly varying measurand. A wirewound potentiometer (a position-sensing device, discussed in Chapter 8) may have a step type of output. Many digital instruments contain an analog-to-digital converter (discussed in Chapter 4), which places well-defined limits on resolution. Most modem instruments have a digital output display, which will result in a resolution error. If, for example, the digital output device has a reading of 1.372, the reading resolution is simply a value of 1 in the last (least significant) digit.

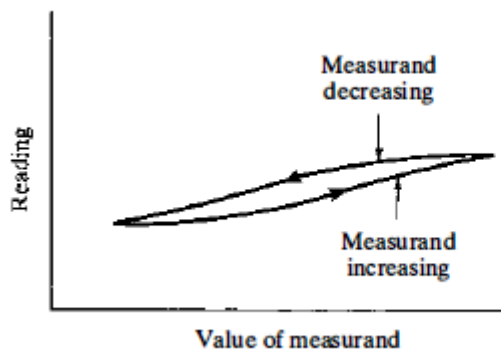


FIGURE 2.5
Effect of hysteresis on instrument reading.

If the device rounds off values, the resolution uncertainty will be in the least significant digit.

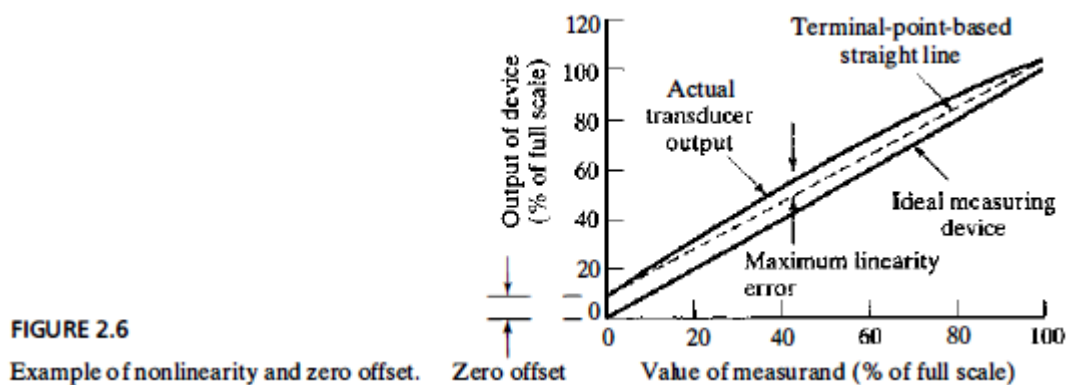
In instruments in which the output is read by comparing a pointer to a scale, the ability to resolve a value of the measurand is limited by a characteristic called the scale **readability**. For example, the thermometer shown in Figure 2.2 has a tick mark every degree. One might think that there may be a maximum uncertainty in reading the thermometer of ± 0.5 degree (reading to the nearest tick mark). However, the human eye can visually interpolate in the interval between the tick marks—perhaps breaking the interval into five parts. The error due to readability may thus be only ± 0.2 degree. Regardless of the spacing of the ticks, the human eye will find it difficult to discriminate differences of less than 0.01 in. (Sweeney, 1953). The manufacturer of the measuring system may well take the output resolution or readability into account in designing the device—the resolution may, in fact, reflect the accuracy of the device.

It is pointless to be able to resolve a reading to an interval that is smaller than the uncertainty interval of the measurement.

Repeatability is the ability of a device to produce the same output reading when the same measurand is applied using the same procedure. Inability to repeat a measurement exactly, a random error known as **repeatability error**, is usually a component of the manufacturers' specification of instrument accuracy. It should be noted that hysteresis is not a cause of repeatability error - hysteresis is a separate error. The concept of random error of a measurement is more general than measuring device repeatability and may include variable factors in the measurement process not caused by the measurement device, such as variation of uncontrolled parameters.

Although not a requirement for a measurement system, it is highly desirable that it have a linear relationship between input and output, as shown in Figure 2.6. This means that the change in output is proportional to the change in the value of the measurand. A linear response is particularly useful since it simplifies the process of calibration, or checking that the instrument has low error. If it is known that the sensor is basically linear and has good precision, only two points in the span need to be checked. A highly nonlinear device must be calibrated at several points. Deviation from true linearity when linearity is assumed is a systematic error called **linearity error**.

Linearity error is usually a component of the specification of accuracy. There are a number of ways to determine a linearity specification, and these are presented in ANSI/ISA (1979). In the method that determines terminal-based linearity, a line is drawn connecting the output values at the extreme ends of the span, as shown in Figure 2.6. The linearity error is the maximum deviation between the straight line and the device output. It is normally presented in the form of a percentage of range or a percentage of span.



Measurement systems normally have a point in their range called the **zero or null point**. For example, a weight scale should read zero pounds when there is no weight on the platform. Most instruments have some kind of mechanism to adjust the zero, and any error in zero adjustment will affect all measurements made using the device. If the device does not have correct output at the zero point; it is said to have a **zero offset**, as shown in Figure 2.6. Furthermore, if the zero offset is not accounted for in using the device, the offset will result in a systematic error at all readings called a **zero error**. In some cases, the null point does not correspond to a zero value for either the measurand or the output, but is simply a measurand value to which the device should initially be adjusted. Manufacturers' specifications of

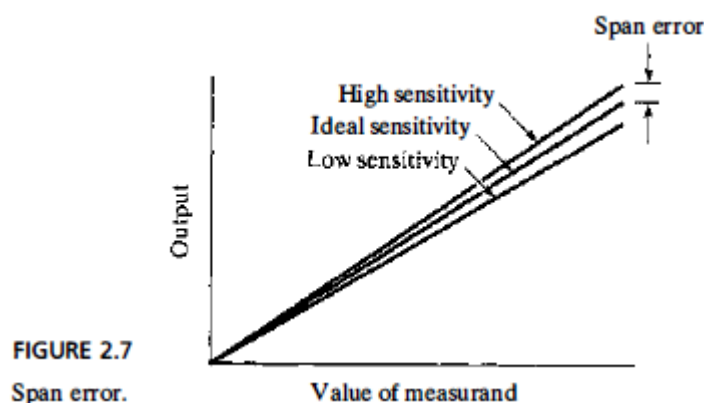
accuracy usually assume that the zero has been adjusted properly. Manufacturers may also specify the maximum expected zero error, often using the term **zero balance**. It is normal to check the zero prior to using a device, and large changes in zero may indicate that the device has been damaged or is malfunctioning.

An important characteristic of a measuring system is the sensitivity, defined as the ratio of the change in magnitude of the output to the change in magnitude of the measurand:

$$\text{sensitivity} = \frac{d(\text{output})}{d(\text{input})} \approx \frac{\Delta \text{output}}{\Delta \text{input}} \quad (2.1)$$

For a thermometer, this could be the change in the height of the mercury column in the stem per degree of temperature change. In mechanical measuring devices, sensitivity is an important and limiting characteristic. In systems with electrical output sensors, the sensitivity can normally be increased using simple amplifiers. However, other limits, such as signal-to-noise ratio, may then become important. In linear systems, the sensitivity is constant throughout the range and is given the symbol **K**. Sensitivity is determined during the process of calibration, and an error in determining the actual sensitivity results in a systematic error called **sensitivity error**, which affects all readings. An error in sensitivity will affect the span, as shown in Figure 2.7. If the span is not as specified, a span **error** will result.

Over a period of time, the output of a measuring system for a fixed measurand may change even though all environmental factors remain constant. This undesirable characteristic is known as **drift**. Many measuring systems are also sensitive to the environmental temperature, a characteristic known as the **thermal stability** of the device. Both drift and thermal stability can affect a number of characteristics of the measuring system and cause additional errors in zero, linearity, hysteresis, and sensitivity. These drift- and thermal-stability-caused errors are not usually included in manufacturers' specifications of accuracy. However, manufacturers often give additional information on drift and **thermal stability** of instruments, which can be used to estimate random uncertainty.



Example 2.2.

An angular-velocity measuring device (tachometer) can measure the speed of a mechanical-shaft in the range 0 to 5000 rpm. It has an accuracy of $\pm 5\%$ of full scale. You notice that when the shaft speed is zero, the device has a reading of 200 rpm. What is the maximum error that you might estimate in reading a shaft speed of 3500 rpm?

Solution: The accuracy specification indicates an uncertainty of $\pm 0.05 \times 5000 = \pm 250$ rpm. Thus all readings have at least this uncertainty. However, there is a zero offset of 200 rpm. This error is in addition to the accuracy uncertainty. Thus the reading might be as much as $250 + 200 = 450$ rpm high. The error estimate can be reduced if the zero is adjusted or if the data are corrected (by subtracting 200 rpm from the reading). If it has been some time since the instrument was calibrated, there might be an additional error in the sensitivity, but this cannot be determined with the information given.

Comment: Accuracy is an uncertainty specification that usually combines unavoidable systematic and random errors associated with using a measuring device. It typically includes hysteresis, linearity, and repeatability error components. It usually does not include other errors, such as zero, drift, and thermal stability. Errors of these types must be considered separately.

2.2.2. Calibration of Measurement Systems

At some point, all measurement systems should undergo **calibration**, a process in which a set of measurements are made of measurand values that can be determined independently. The readings can then be compared to the known "true" values and the errors determined. The number of different values of measurand required for the calibration process depends on the type of measurement system and the application. In some calibration processes, the value of a measurand is known because it is a **standard**. In other calibration processes, another measurement system of known accuracy can be used to determine the value of the measurand. The use of standards is the more reliable approach, but the latter approach is often more practical.

Standards for Calibration Standards of measurement have been important to commerce for a very long time since it is important to the purchaser to know that the weight or length of a purchase is accurate. The relatively recent expansion in science dramatically increased the need for standards. International conferences have established **primary standards** defining the units for seven physical variables. These units are known as the International System (SI) base units (Table 2.1).

TABLE 2.1 SI Base Units	
Physical variable	SI unit
Mass	kg
Time	second
Length	meter
Temperature	kelvin
Electric current	ampere
Amount of a substance (mole)	mole
Light intensity	candela

The standard for mass is the International Prototype Kilogram, which is a platinum-iridium cylinder kept at the International Bureau of Weights and Measures in France.

The standard for time, the second, has been defined as: "the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom" (Wildhack and Jennings, 1992). Although this may seem obscure, it is a standard reproducible in any properly equipped laboratory.

The standard for length, the standard meter, is defined as "the length of the path traveled by light in a vacuum during a time of 11299,792,458 of a second" (Wildhack and Jennings, 1992).

The standard for temperature is more complicated than for the other base units since it must be specified over a wide range of values. The standard is known as the International Temperature Scale of 1990 (ITS-90) (Preston-Thomas, 1990). The measure of temperature is the thermodynamic temperature, and the unit is the kelvin, defined as 1/273.16 of the thermodynamic temperature of the triple point of water, that temperature where solid, liquid, and vapor phases of pure water exist together in thermal equilibrium. The standard temperature scale extends from 0.65 K to the highest temperatures that can be determined by measuring thermal radiation. While the details of the complete standard are beyond the scope of the present book, in the range of greatest interest to engineer, the standard is fairly straightforward. From the triple point of hydrogen, 13.8033 K, to the freezing point of silver, 961.78 K, the scale is defined by means of a platinum resistance thermometer (see Chapter 9), which is calibrated at a set of fixed points shown in Table 2.2. Above the freezing point of silver, the temperature standard is based on a relationship between the thermal radiation from an object at the measured temperature to the thermal radiation from an object at the temperature of freezing silver, gold, or copper. Guidelines for using the ITS-90 standard are given by Mangum and Furukawa (1990).

The standard for electric current, the ampere, is "that constant current which, if maintained in two straight parallel conductors of infinite length and of negligible circular sections, and placed 1 meter apart in a vacuum would produce a force equal to 2×10^{-7} newton per meter of length" (Wildhack and Jennings, 1992).

The standard for the mole is "the amount of a substance of a system which contains as many elementary entities as there are atoms in 0.012 kg of carbon-12" (Wildhack and Jennings, 1992).

TABLE 2.2 Fixed Points of International Temperature Scale of 1990 (ITS-90)

Fixed point ^a	Temperature (K)
Triple point of hydrogen	13.8033
Triple point of neon	24.5561
Triple point of oxygen	54.3584
Triple point of argon	83.8058
Triple point of mercury	234.3156
Triple point of water	273.16
Melting point of gallium	302.9146
Freezing point of lanthanum	429.7485
Freezing point of tin	505.078
Freezing point of zinc	692.677
Freezing point of aluminum	933.473
Freezing point of silver	1234.93
Freezing point of gold	1337.33
Freezing point of copper	1357.77

Source: Preston-Thomas (1990).

^aMelting and freezing points are at a pressure of 101.325 Pa.

The standard for *light intensity*, the candela, is "the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} hertz and of which the radiant intensity in that direction is 1/683 watt per steradian" (Wildhack and Jennings, 1992).

All the primary standards except mass can, in theory, be reproduced in any laboratory having the proper equipment. They are called **reproducible standards**. Mass, the only standard that has an exact physical location, is called a **fixed standard**. Although all primary standards except mass can, in principle, be reproduced in any laboratory, this is usually not practical, and the primary standards are difficult to apply for every-day calibration activities. As a result, it is normal to create **secondary standards**. These standards must be traceable to the primary standards—that is, at some point, the secondary standard was compared either to a primary standard or to another secondary standard that is traceable to a primary standard. Laboratories can have masses, accurately sized pieces of metal called gage blocks, quartz crystal clocks, and other practical secondary standards for calibration purposes.

The foregoing list of base primary standards does not include most of the common variables for which calibration is required (force and voltage are obvious examples). This is because standards for all other physical variables are derived from variables having the base primary standards by using physical laws or scientific definitions. The standard for force, for example, is defined by Newton's second law, $F = ma$. Since primary standards have been established for mass, length, and time, the standard for force can be computed. Laboratory secondary standards can also be created for these other variables—a standard battery for voltage is an example. Standard names and definitions are provided for virtually all derived units, and these are documented in sources such as the International Organization for Standardization (ISO, 1979).

Some calibration processes are quite difficult, and it is not practical for all laboratories to keep all necessary standards. Laboratories may purchase instruments that have been calibrated by the manufacturers. If it is expected that the uncertainty of a measuring device will change with time, it may have to be returned to the manufacturer (or another laboratory) periodically for calibration. Most laboratories can afford to have some secondary working standards. For example, standards for mass and length can be purchased at moderate cost and, with reasonable care, are very durable.

Static Calibration Process In the calibration of a measuring system, the system is used to make measurements of known values of the measurand. Various values of the measurand must be used to cover the intended range of use of the measurement system. As mentioned, it is best that the values of the measurand be known because they are standards, while from a practical standpoint, however, it is often easier to determine the values of some measurands using another measuring system of known accuracy. It is preferable that this second measuring system be based on simple physical principles so that the experimenter will have confidence that it maintains its accuracy. For example, a balance scale for measuring force is more likely to maintain its accuracy than a scale that uses springs.

The first step in the calibration process is to take a set of data consisting of measurement system output as a function of the measurand. A graphical presentation of these data is known as the **calibration curve**. Although this curve can be used directly to interpret the instrument output, in most cases the data are correlated using a mathematical function (the process of curve fitting). In some cases, the calibration data will be used to confirm an existing function. The correlating function may be a straight line, a parabola, a higher-order polynomial, or a more complicated function. This function is used to obtain values of the measurand from the measuring device output in the actual experiment. Finally, an estimate should be made of the calibrated system's contribution to the overall experimental uncertainty. An estimate should

be made of residual systematic error, and, depending on the expected use of the measurement system, it may also be necessary to estimate the expected random error.

The document ANSI/ISA (1979) describes a standardized procedure for static calibration that is used by many instrument manufacturers. In the first step of this procedure, the calibration data are taken by applying known values of the measurand to the measuring system. The measurand values are incrementally increased from the low end of the range to the top end of the range. The measurand values are then incrementally reduced to the low end of the range. This process is repeated several times (cycles). The procedure then describes how to determine specific values for such errors as linearity, hysteresis, and repeatability as well as overall accuracy. In this procedure, the accuracy statement effectively combines the errors due to nonlinearity, hysteresis, and nonrepeatability. The term static indicates that the process is performed in a manner in which time is not a factor-the measurand is changed slowly, and the device is allowed to come to equilibrium prior to taking a reading.

The sequential calibration process just described does not, in general, duplicate the actual measurement process. In a typical measurement situation, the experimenter may not be able to determine if the measurand is increasing or decreasing prior to the measurement. A calibration process in which the values of the measurands are randomly selected is often used. With a random calibration, it may be difficult to separate hysteresis errors from repeatability errors, and hysteresis may have to be treated as part of the random error. This limitation may not be important.

Example 2.3, based on information in ANSI/ISA (1979), is used to demonstrate the static calibration process to determine instrument accuracy and other errors. This process will not determine thermal stability and drift characteristics. Furthermore, it will not account for errors associated with the application, such as spatial errors, nor will it account for dynamic (time-varying) effects. Additional calibration procedures may be required for specific experiments. As discussed in Chapter 7, the error characteristics determined by the method shown in Example 2.3 are in a usable but not ideal form for detailed uncertainty analysis.

Example 2.3. Calibration of a Weighing Scale

A low-cost, nominally 0 to 5-lb spring weighing scale [Figure E2.3(a)] has been calibrated by placing accurate weights on its platform. The values of the applied weights range from 0 to 5 lb in 0.5-lb increments. The weights are applied in a sequential manner, starting at the lowest value, increasing to the largest value (up data) and then decreasing to the lowest value (down data). Five such cycles were performed, and the results of the measurements are presented in Table E2.3(a). As suggested in ANSI/ISA (1979), several cycles were completed before the data recording started. The data recording then started in the middle of the up portion of cycle 1 and ended in the up portion of cycle 6, giving five complete cycles.

Fit a straight line to the data and determine the accuracy, hysteresis, and linearity errors. Also, make estimates of the maximum systematic and random errors.

Solution: The data of Table E2.3(a), which have been plotted in Figure E2.3(b), fall into two bands. The division into two bands is caused by hysteresis in the system-the lower band is for increasing measurand and the upper band for decreasing measurand.

There are a number of methods used to determine a suitable straight-line fit to a set of instrument data. A technique called the method of least squares (linear regression) will be

described in Chapter 6. At this point, it is convenient simply to "eyeball" a line through the data to minimize the maximum deviations of the data from the line. This process will approximate a least-squares linear fit. The resulting correlating equation takes the form

$$R = 1.290W - 0.374$$

where R is the reading of the scale and W is the actual weight. Alternatively, we might have given the equation in the form of the weight versus the reading ($W = 0.775R + 0.290$).

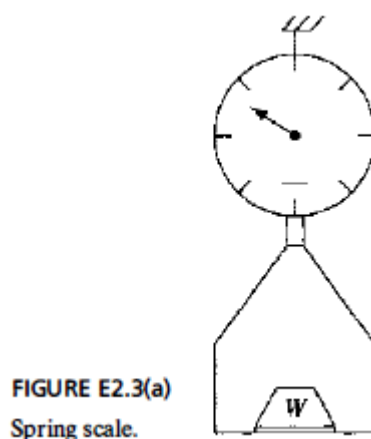


TABLE E2.3(a) Scale Calibration Data

True weight (lb)	Scale reading					
	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6
0.5		0.2	0.08	0.17	0.19	0.11
1		0.7	0.78	0.64	0.61	0.7
1.5		1.18	1.26	1.25	1.24	1.23
2		1.81	1.93	1.81	1.93	1.88
2.5	2.62	2.49	2.46	2.46	2.58	2.53
3	3.15	3.18	3.24	3.28	3.13	
3.5	3.9	3.84	3.86	3.97	3.96	
4	4.59	4.71	4.61	4.6	4.6	
4.5	5.41	5.35	5.49	5.46	5.39	
5	6.24	6.27	6.1	6.24	6.16	
4.5	5.71	5.74	5.78	5.87	5.82	
4	4.96	5.11	5.08	5.03	5.03	
3.5	4.22	4.34	4.21	4.22	4.24	
3	3.57	3.64	3.66	3.55	3.67	
2.5	2.98	2.86	2.98	2.98	2.94	
2	2.22	2.23	2.26	2.29	2.26	
1.5	1.57	1.7	1.69	1.63	1.57	
1	1.07	1.07	1.11	1.16	1.11	
0.5	0.52	0.61	0.61	0.61	0.45	
0	0.02	0.08	0.08	-0.03	0.06	

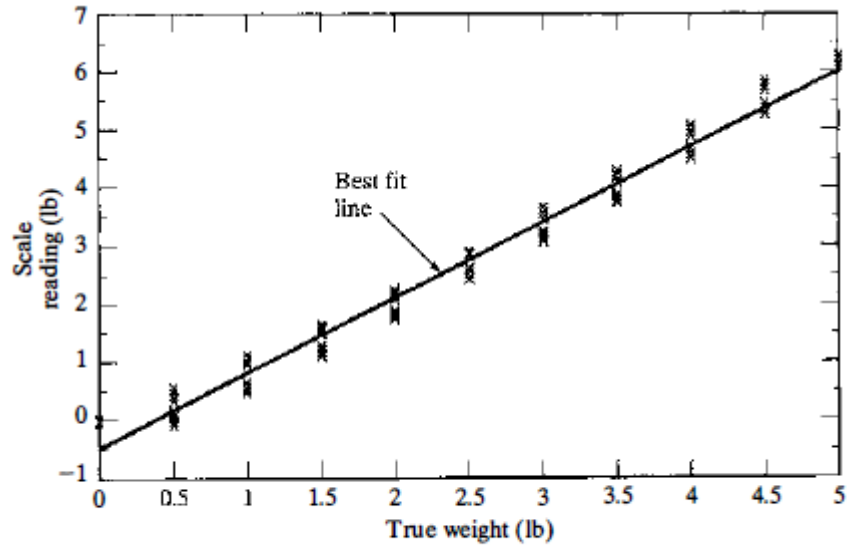


FIGURE E2.3(b)
Plot of scale readings.

Although Figure E2.3(b) is suitable for determining the correlating function, to evaluate the errors it is best to present the data in the form of what is called a deviation plot. For each measured value, the difference (deviation) between the measured value and the best-fit equation is evaluated. These values are presented in Table E2.3(b). The last three columns of the table will be discussed later. The deviation data are plotted in Figure E2.3(c). A deviation value of zero indicates that a calibration data point has a value exactly as predicted by the correlating function.

TABLE E2.3(b) Scale Deviation Data

True weight (lb)	Deviation						Average of cycles	Average up-down	Repeat
	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6			
0								0.41	
0.5		-0.07	-0.19	-0.1	-0.08	-0.16	-0.12	0.085	0.12
1		-0.22	-0.14	-0.28	-0.31	-0.22	-0.23	-0.025	0.17
1.5		-0.38	-0.3	-0.31	-0.32	-0.33	-0.33	-0.13	0.08
2		-0.4	-0.28	-0.4	-0.28	-0.33	-0.34	-0.15	0.12
2.5	-0.23	-0.36	-0.39	-0.39	-0.27	-0.32	-0.35	-0.125	0.16
3	-0.35	-0.32	-0.26	-0.22	-0.37		-0.3	-0.09	0.15
3.5	-0.24	-0.3	-0.28	-0.17	-0.18		-0.23	-0.06	0.13
4	-0.2	-0.08	-0.18	-0.19	-0.19		-0.17	0.04	0.12
4.5	-0.02	-0.08	0.06	0.03	-0.04		-0.01	0.17	0.14
5	0.16	0.19	0.02	0.16	0.08		0.12	0.12	0.17
4.5	0.28	0.31	0.35	0.44	0.39		0.35		0.16
4	0.17	0.32	0.29	0.24	0.24		0.25		0.15
3.5	0.08	0.2	0.07	0.08	0.1		0.11		0.13
3	0.07	0.14	0.16	0.05	0.17		0.12		0.12
2.5	0.13	0.01	0.13	0.13	0.09		0.1		0.12
2	0.01	0.02	0.05	0.08	0.05		0.04		0.07
1.5	0.01	0.14	0.13	0.07	0.01		0.07		0.13
1	0.15	0.15	0.19	0.24	0.19		0.18		0.09
0.5	0.25	0.34	0.34	0.34	0.18		0.29		0.16
0	0.39	0.45	0.45	0.34	0.43		0.41		0.11

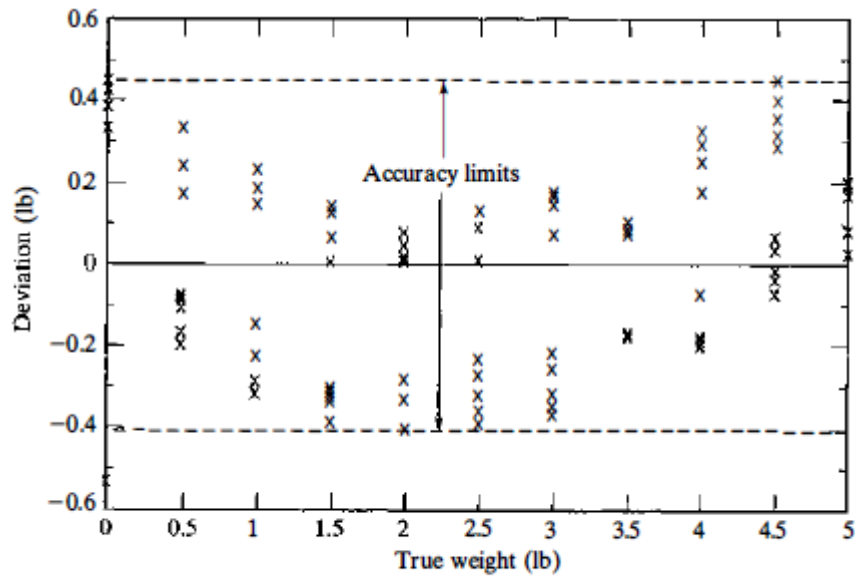


FIGURE E2.3(c)

Plot of deviation data for scale calibration.

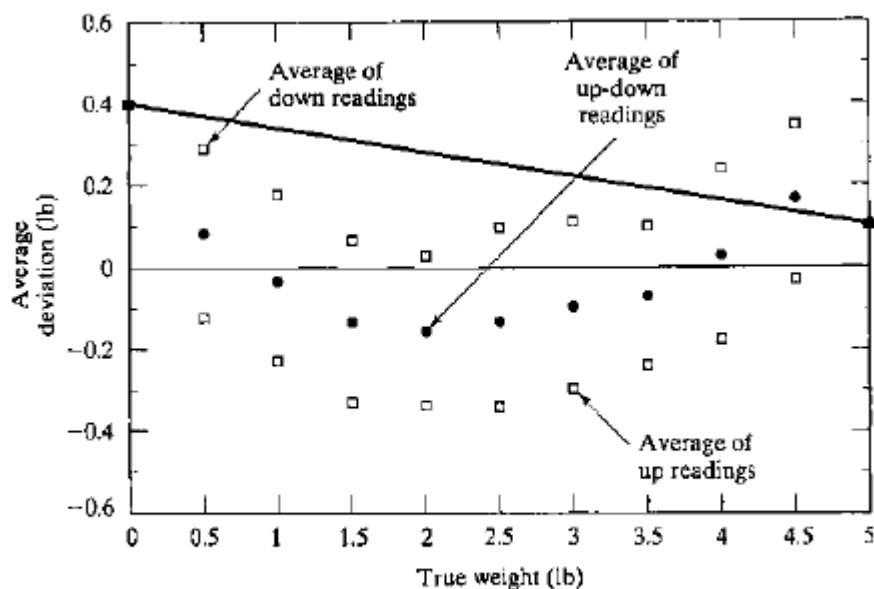


FIGURE E2.3(d)

Average deviation data for scale calibration.

Using Figure E2.3(c), we find that it is a simple matter to estimate the accuracy of this weighting device. Two lines are drawn, parallel to the horizontal axis, such that all the data are contained between the two lines. For this device, the accuracy limits are +0.45 lb and -0.40 lb. Accuracy is frequently presented as a percent of output span. In this case, the output, as predicted by the curve fit, varies from -0.374 to +6.076, giving a span of $6.076 - (-0.374) = 6.45$ lb. Accuracy then becomes +7.0% and -6.2% of the output span. It should be noted that accuracy is a bounding error statement—it includes all residual errors that will occur when the instrument is used with the given linear fit to the data. Accuracy does not include errors due to drift or thermal-stability effects, which require additional calibration procedures.

To evaluate the nonlinearity of the data, it is necessary to average all the data taken for each value of the weight. Not only are the values for all the cycles at a given weight averaged, but the "up" and the "down" values are also averaged. These results are presented in the column labeled "Average up-down" in Table E2.3(b). These data are plotted with the solid symbols in Figure E2.3(d). A line has been drawn connecting the terminal points of the data, and the terminal linearity can be evaluated as 0.44 lb, which is about 6.9% of output span. There are other measures of linearity, as described in ANSI/ISA (1979) and Norton (1982).

The repeatability error is the maximum variability of successive measurements of the same value of input approached from the same direction. The column of Table E2.3(b) labeled "Repeat" is this variability for each of the test conditions. The maximum variation, 0.17 lb, occurs at an "up" reading of 1 lb. This corresponds to 2.6% (or $\pm 1.3\%$) of output span.

ANSI/ISA (1979) indicates that the hysteresis error is evaluated as the maximum difference between the "up" and the corresponding "down" reading for any of the calibration cycles. Evaluating this difference for all the data in Table E2.3(b), the maximum difference is found to be 0.52 lb and occurs in cycles 3 and 4 at 2.5lb. This becomes 8.1% (or $\pm 4.05\%$) of output span. Technically, this is not simply the hysteresis error but a combination of the hysteresis error plus another error called dead band. These two errors are frequently combined (and often simply called hysteresis error). The interested reader can find a description of the distinction between these two errors in ANSI/ISA (1979).

At this point, it is useful to estimate the total uncertainty due to systematic error and that due to random error, a process not considered in ANSI/ISA (1979). Systematic error is the error that can be expected if a large number of readings are taken of a particular value of the measurand using the same procedure and then averaged. Table E2.3(b) shows these results in the column labeled "Average of cycles," and the data are plotted in Figure E2.3(d). Note that the "up" and "down" values are treated as distinct data. An estimate of the maximum systematic error would be the maximum deviations of these averaged calibration data. The resulting error limits are +0.41 lb and -0.35 lb, which become +6.4% and -5.4% of output span. Systematic errors are discussed further in Chapter 7.

Statistical techniques provide the best approach to evaluating random uncertainty. However, an estimate of the random uncertainty is the repeatability error. The repeatability error limits are $\pm 1.3\%$ (see above). Better techniques to evaluate the random error are developed in Chapters 6 and 7. In addition, factors not considered in this calibration procedure (such as uncontrolled changes in ambient temperature) can also contribute to random uncertainty.

Comments: The accuracy numbers determined by the process described above can be used to estimate the maximum uncertainty in measurements made with the scale. The accuracy values (+0.45/-0.40 lb) apply to the device output when a given weight is placed on the platform. Using the correlating function, it is a simple matter not only to determine the weight corresponding to a given output reading but also the uncertainty in that weight. It is left for the reader to show that for a given reading, the uncertainty in the applied weight will be +0.35/-0.31 lb.

There are techniques by which the uncertainty in weight values determined with this scale might be reduced. First, a nonlinear function such as a parabola might be used to curve-fit the data. Second, the uncertainty interval can be made a function of the reading, resulting in a reduction in the uncertainty interval for some portions of the range. However, hysteresis is a major cause of the poor accuracy, and this cannot be reduced by calibration procedures.

The type of calibration described in Example 2.2 is a **static calibration**. However, many measuring systems are used in situations in which the measurand is changing rapidly. For such situations, a static calibration is inadequate. If possible, the instrument can be subjected to a dynamic calibration. Dynamic calibration is usually a more complicated process and the procedure depends on the device being calibrated and often, the specific application. In many cases, dynamic response is determined entirely by analysis.

2.3. DYNAMIC MEASUREMENTS

If a measurand is unchanging in time and if the measurement system instantaneously shows an equilibrium response to the measurand, the measurement process is said to be static. In the general case, however, when the measurand is changing in time and the measuring system does not show instantaneous response, the measurement process is said to be dynamic. For example, the use of an oral thermometer to measure a person's body temperature is a dynamic measurement since the measurement process must be continued for several minutes for the thermometer to come into equilibrium with body temperature. In making dynamic measurements, the experimenter must account for the dynamic characteristics of the measuring system, the dynamic interaction between the measuring system and the test system, and the dynamic characteristics of the test system.

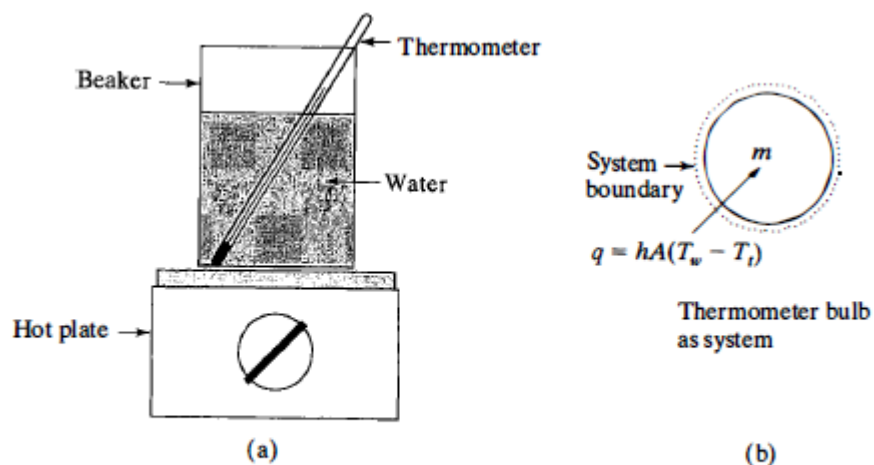


FIGURE 2.8
Dynamic temperature measurement.

When a measurement is dynamic, there is usually an error introduced into the measurement, and the experimenter must take actions to minimize this error. For example, consider the measurement of the temperature of the water in a beaker that is being heated on a hot plate as shown in Figure 2.8(a). Considering the bulb of the thermometer as a system [Figure 2.8(b)], we can apply the first law of thermodynamics. Since there is no work term, we obtain

$$q = hA(T_w - T_t) = mc \frac{dT_t}{dt}$$

or

$$(T_w - T_t) = \frac{mc}{hA} \frac{dT_t}{dt} \quad (2.2)$$

where Newton's law of cooling [$q = hA(T_w - T_t)$] has been used to estimate q , and the symbols are defined as follows:

A	bulb surface area
T_w	water temperature
T_t	thermometer temperature
h	heat transfer coefficient
m	bulb mass
c	bulb specific heat
dT_t/dt	time rate of change of the water temperature

There would be no error in the measurement if $T_w = T_t$. However, if the water is being heated, dT_t/dt is nonzero, so $T_w - T_t$ is nonzero and there exists an inherent dynamic measurement error. The term mc/hA , called the **time constant** of the system, has dimensions of time and is usually given the symbol τ . To minimize the dynamic error for this situation, τ should be made as small as is practical.

The dynamic response of a measurement system can usually be placed into one of three categories: zero order, first order, and second order. These categories are based on the order of the differential equation needed to describe the dynamic response. Ideal **zero-order** systems respond instantly to measurands, although no instrument is truly zero order. However, the dynamic response of many instruments approximate zero-order behavior when measuring slowly changing measurands. **First-order** measurement systems show capacitance-type energy storage effects. Mechanical analogs to capacitance are springs and devices that store thermal energy. The common thermometer discussed above is an example of a first-order system. **Second-order** systems have inertial effects of inductance or accelerated mass as well as capacitance energy storage. Common spring-mass systems are second order - the mechanical bathroom scale is an example. Second-order systems include a characteristic called damping, which dissipates energy. Second-order systems with low damping are called **underdamped** and can show oscillatory response. Highly damped second-order systems are called **overdamped** and will not show oscillatory behavior. The level of damping that divides these two modes of response is called **critical damping**.

There are specifications for the dynamic response of an instrument that can be used to select an appropriate device. The two most common specifications characterize the response of the measurement system to a step change in input (called **transient response**) and the response to a range of sinusoidal inputs (called **frequency response**).

A step change in input to a measuring system is shown in Figure 2.9(a), and some typical system responses are shown in Figure 2.9(b). Y is the change in the device output, and Y_e is

the equilibrium change in the output of the system, the output that will occur after significant time has passed. Response A is typical of first-order and overdamped second-order devices. Response B is typical of underdamped second-order systems.

First-order systems and overdamped second-order systems show an asymptotic response. For first-order systems, this takes an exponential form:

$$\frac{y}{y_e} = 1 - e^{-t/\tau} \quad (2.3)$$

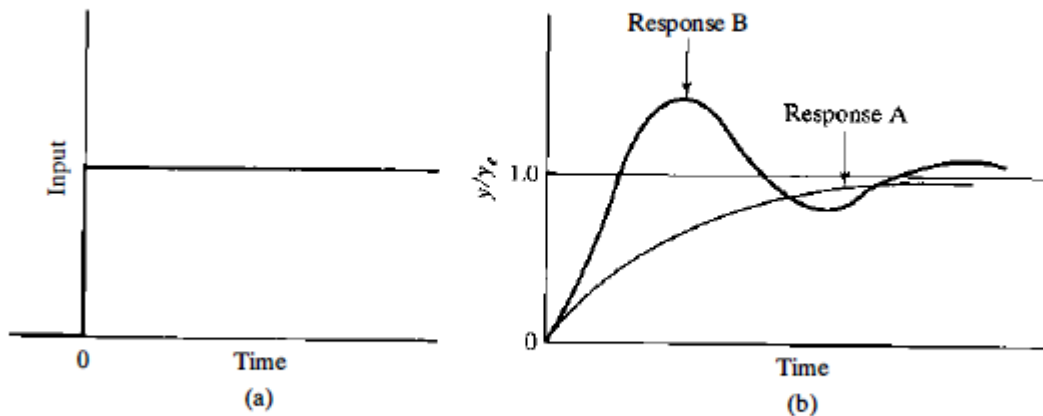


FIGURE 2.9

Application of a step change in measurand; (a) time variation of measurand; (b) typical types of system responses.

The time constant, τ , thus determines the curve and is a useful numerical specification of the transient response of the instrument. It is the time at which the response, y/y_e , has a value of $1 - 1/e = 0.632$. Overdamped second-order systems show a response which is similar to, but somewhat more complicated than, that predicted by Eq. (2.3). As a result, the concept of time constant is not exactly applicable to second-order systems. On the other hand, **response time** is a more general term used to specify transient response and can be used for first- and second-order systems. For example, a 95 % response time is the time at which y/y_e has a value of 0.95. Another measure of transient response is the **rise time**. This is usually the time that it takes y/y_e to increase from 0.1 to 0.9.

For oscillatory responses [response B in Figure 2.9(b)], another term, called **settling time**, is more useful. This is the time until the amplitude of the oscillations are less than some fraction (e.g., 10%) of the equilibrium response.

Generally speaking, for minimal dynamic error, the appropriate measure of transient response (e.g., response time or settling time) should be small compared to the expected time for the measurand to change in an experiment. If the measurand changes in a time comparable to the transient response time, there will be a significant dynamic measurement error.

Frequency response is another useful measure of dynamic response. The output of the system is determined for a pure sine-wave input. This process is repeated for a range of frequencies to produce a frequency-response curve such as that shown in Figure 2.10. As shown in Figure 2.10, there is the range of sinusoidal input frequencies over which the measuring system gives a constant ratio of output amplitude to input amplitude.

This range of frequencies is called the **bandwidth** of the device. If the device is used for measurements outside the bandwidth, the ratio of output to input amplitudes will change and significant systematic errors are likely to occur. Most input signals are not sinusoidal and are rather complicated functions of time. It is often necessary to decompose complicated waveforms into sinusoidal components, and these components can be used to determine the required frequency response of the measuring system.

In measuring frequency response, it will usually be noted that there is a phase difference between the input and the output. This phase difference will usually depend on the frequency and must be considered in many situations.

In some cases, the frequency response of a second-order system is specified with a single number called the **natural frequency**. In the spring scale example, if a weight is simply dropped on the platform, the device will oscillate for a period of time at the natural frequency. In most cases, dynamic response will be degraded if the measurand varies at a frequency greater than 0.2 to 0.4 times the natural frequency.

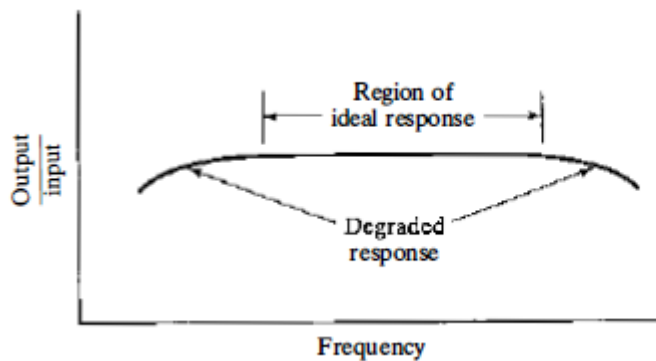


FIGURE 2.10
System frequency response.

There are, however, some devices that operate correctly only at frequencies greater than the natural frequency.

Manufacturers' specifications of dynamic response are usually available and can be used for many components of measuring systems. Unfortunately, for some sensing elements and transducers, it is not possible for the manufacturer to supply all the needed dynamic specifications. This is because the dynamic response of the transducer depends not only on the characteristics of the transducer itself but also on the manner in which the transducer is used. For example, the rise time of a thermometer used to measure the temperature of a fluid depends on the properties and motion of the fluid as well as the thermometer itself.