Methods of measurement and codes

MED 371

By Dr.Gomaa A. Abdelrazek

2026-2027

Brief Contents

PART I: ENGINEERING METROLOGY 1

1. Basic Principles of Engineering Metrology
2. Standards of Measurement
3. Limits, Fits, and Tolerances
4. Linear Measurement
5. Angular Measurement
6. Comparators
7. Optical Measurement and Interferometry
8. Metrology of Gears and Screw Threads
9. Metrology of Surface Finish
10. Miscellaneous Metrology
11. Inspection and Quality Control

PART II: MECHANICAL MEASUREMENTS

12. Measurement Systems
13. Transducers
14. Measurement of Force, Torque, and Strain
15. Measurement of Temperature
16. Pressure Measurements

CHAPTER-1

Basic Principles of Engineering Metrology

After studying this chapter, the reader will be able to

- understand the importance of metrology
- appreciate the significance of inspection
- appreciate the concepts of accuracy and precision
- explain the objectives of metrology and measurements
- understand the general measurement concepts
- elucidate the different sources and types of errors
- compare the different types of measurements

1.1 INTRODUCTION

The importance of metrology as a scientific discipline gained momentum during the industrial revolution. Continuing technological advancement further necessitated refinement in this segment. Metrology is practised almost every day, often unknowingly, in our day-to-day tasks. Measurement is closely associated with all the activities pertaining to scientific, industrial, commercial, and human aspects. Its role is ever increasing and encompasses different fields such as communications, energy, medical sciences, food sciences, environment, trade, transportation, and military applications. Metrology concerns itself with the study of measurements. It is of utmost importance to measure different types of parameters or physical variables and quantify each of them with a specific unit. Thus, measurement is an act of assigning an accurate and precise value to a physical variable. The physical variable then gets transformed into a measured variable. Meaningful measurements require common measurement standards and must be performed using them. The common methods of measurement are based on the development of international specification standards. These provide appropriate definitions of parameters and protocols that enable standard measurements to be made and also establish a common basis for comparing measured values. In addition, metrology is also concerned with the reproduction, conservation, and transfer of units of measurements and their standards. Measurements provide a basis for judgements about process information, quality assurance, and process control.

Design is one of the major aspects of all branches of engineering. A product/system comprising several elements has to be properly designed to perform the required (desired) function. In order to test whether functioning of the elements constituting the product/system meets the design expectation, and to finally assess the functioning of the whole system, measurements

2. To ensure uniformity of measurements
3. To carry out process capability studies to achieve better component tolerances
4. To assess the adequacy of measuring instrument capabilities to carry out their respective measurements
5. To ensure cost-effective inspection and optimal use of available facilities
6. To adopt quality control techniques to minimize scrap rate and rework
7. To establish inspection procedures from the design stage itself, so that the measuring methods are standardized
8. To calibrate measuring instruments regularly in order to maintain accuracy in measurement
9. To resolve the measurement problems that might arise in the shop floor
10. To design gauges and special fixtures required to carry out inspection
11. To investigate and eliminate different sources of measuring errors

1.6 GENERAL MEASUREMENT CONCEPTS

We know that the primary objective of measurement in industrial inspection is to determine the quality of the component manufactured. Different quality requirements, such as permissible tolerance limits, form, surface finish, size, and flatness, have to be considered to check the conformity of the component to the quality specifications. In order to realize this, quantitative information of a physical object or process has to be acquired by comparison with a reference. The three basic elements of measurements (schematically shown in Fig. 1.3), which are of significance, are the following:

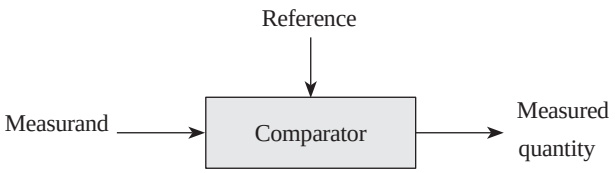


Fig. 1.3 Elements of measurement

1. *Measurand*, a physical quantity such as length, weight, and angle to be measured
2. *Comparator*, to compare the measurand (physical quantity) with a known standard (reference) for evaluation
3. *Reference*, the physical quantity or property to which quantitative comparisons are to be made, which is internationally accepted

All these three elements would be considered to explain the direct measurement using a calibrated fixed reference. In order to determine the length (a physical quantity called measurand) of the component, measurement is carried out by comparing it with a steel scale (a known standard).

1.6.1 Calibration of Measuring Instruments

It is essential that the equipment/instrument used to measure a given physical quantity is validated. The process of validation of the measurements to ascertain whether the given physical quantity conforms to the original/national standard of measurement is known as *traceability* of the standard. One of the principal objectives of metrology and measurements is to analyse the uncertainty of individual measurements, the efforts made to validate each measurement with a given equipment/instrument, and the data obtained from it. It is essential

that traceability (which is often performed by a calibration laboratory having conformity with a proven quality system with such standards) should disseminate to the consumers. Calibration is a means of achieving traceability. One of the essential aspects of metrology is that the results of measurements obtained should be meaningful. To accomplish this, calibration of any measuring system/instrument is very essential. Calibration is the procedure used to establish a relationship between the values of the quantities indicated by the measuring instrument and the corresponding values realized by standards under specified conditions. It refers to the process of establishing the characteristic relationship between the values of the physical quantity applied to the instrument and the corresponding positions of the index, or creating a chart of quantities being measured versus readings of the instrument. If the instrument has an arbitrary scale, the indication has to be multiplied by a factor to obtain the nominal value of the quantity measured, which is referred to as *scale factor*. If the values of the variable involved remain constant (not time dependent) while calibrating a given instrument, this type of calibration is known as *static calibration*, whereas if the value is time dependent or time-based information is required, it is called *dynamic calibration*. The relationship between an input of known dynamic behaviour and the measurement system output is determined by dynamic calibration.

The main objective of all calibration activities is to ensure that the measuring instrument will function to realize its accuracy objectives. General calibration requirements of the measuring systems are as follows: (a) accepting calibration of the new system, (b) ensuring traceability of standards for the unit of measurement under consideration, and (c) carrying out calibration of measurement periodically, depending on the usage or when it is used after storage.

Calibration is achieved by comparing the measuring instrument with the following: (a) a primary standard, (b) a known source of input, and (c) a secondary standard that possesses a higher accuracy than the instrument to be calibrated. During calibration, the dimensions and tolerances of the gauge or accuracy of the measuring instrument is checked by comparing it with a standard instrument or gauge of known accuracy. If deviations are detected, suitable adjustments are made in the instrument to ensure an acceptable level of accuracy. The limiting factor of the calibration process is repeatability, because it is the only characteristic error that cannot be calibrated out of the measuring system and hence the overall measurement accuracy is curtailed. Thus, repeatability could also be termed as the minimum uncertainty that exists between a measurand and a standard. Conditions that exist during calibration of the instrument should be similar to the conditions under which actual measurements are made. The standard that is used for calibration purpose should normally be one order of magnitude more accurate than the instrument to be calibrated. When it is intended to achieve greater accuracy, it becomes imperative to know all the sources of errors so that they can be evaluated and controlled.

1.7 ERRORS IN MEASUREMENTS

While performing physical measurements, it is important to note that the measurements obtained are not completely accurate, as they are associated with uncertainty. Thus, in order to analyse the measurement data, we need to understand the nature of errors associated with the measurements.

Therefore, it is imperative to investigate the causes or sources of these errors in measurement

systems and find out ways for their subsequent elimination. Two broad categories of errors in measurement have been identified: systematic and random errors.

1.7.1 Systematic or Controllable Errors

A systematic error is a type of error that deviates by a fixed amount from the true value of measurement. These types of errors are controllable in both their magnitude and their direction, and can be assessed and minimized if efforts are made to analyse them. In order to assess them, it is important to know all the sources of such errors, and if their algebraic sum is significant with respect to the manufacturing tolerance, necessary allowance should be provided to the measured size of the workpiece. Examples of such errors include measurement of length using a metre scale, measurement of current with inaccurately calibrated ammeters, etc. When the systematic errors obtained are minimum, the measurement is said to be extremely accurate. It is difficult to identify systematic errors, and statistical analysis cannot be performed. In addition, systematic errors cannot be eliminated by taking a large number of readings and then averaging them out. These errors are reproducible inaccuracies that are consistently in the same direction. Minimization of systematic errors increases the accuracy of measurement. The following are the reasons for their occurrence:

1. Calibration errors
2. Ambient conditions
3. Deformation of workpiece
4. Avoidable errors

Calibration Errors

A small amount of variation from the nominal value will be present in the actual length standards, as in slip gauges and engraved scales. Inertia of the instrument and its hysteresis effects do not allow the instrument to translate with true fidelity. Hysteresis is defined as the difference between the indications of the measuring instrument when the value of the quantity is measured in both the ascending and descending orders. These variations have positive significance for higher-order accuracy achievement. Calibration curves are used to minimize such variations. Inadequate amplification of the instrument also affects the accuracy.

Ambient Conditions

It is essential to maintain the ambient conditions at internationally accepted values of standard temperature (20 ºC) and pressure (760 mmHg) conditions. A small difference of 10 mmHg can cause errors in the measured size of the component. The most significant ambient condition affecting the accuracy of measurement is temperature. An increase in temperature of 1 ºC results in an increase in the length of C25 steel by 0.3 μm, and this is substantial when precision measurement is required. In order to obtain error-free results, a correction factor for temperature has to be provided. Therefore, in case of measurements using strain gauges, temperature compensation is provided to obtain accurate results. Relative humidity, thermal gradients, vibrations, and CO_2 content of the air affect the refractive index of the atmosphere. Thermal expansion occurs due to heat radiation from different sources such as lights, sunlight, and body temperature of operators.

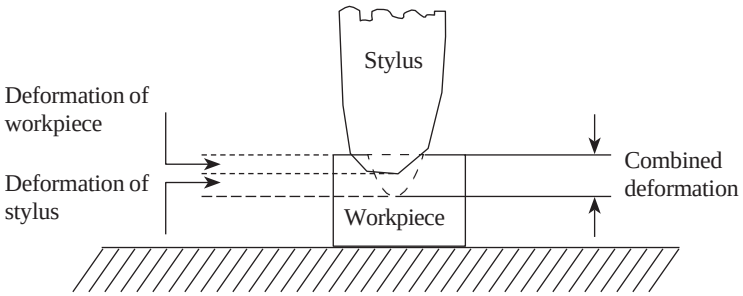


Fig. 1.4 Elastic deformation due to stylus pressure

Deformation of Workpiece

Any elastic body, when subjected to a load, undergoes elastic deformation. The stylus pressure applied during measurement affects the accuracy of measurement. Due to a definite stylus pressure, elastic deformation of the workpiece and deflection of the workpiece shape may occur, as shown in Fig. 1.4. The magnitude of deformation depends on the applied load, area of contact, and mechanical properties of the material of the given workpiece. Therefore, during comparative measurement, one has to ensure that the applied measuring loads are same.

Avoidable Errors

These include the following:

Datum errors Datum error is the difference between the true value of the quantity being measured and the indicated value, with due regard to the sign of each. When the instrument is used under specified conditions and a physical quantity is presented to it for the purpose of verifying the setting, the indication error is referred to as the datum error.

Reading errors These errors occur due to the mistakes committed by the observer while noting down the values of the quantity being measured. Digital readout devices, which are increasingly being used for display purposes, eliminate or minimize most of the reading errors usually made by the observer.

Errors due to parallax effect Parallax errors occur when the sight is not perpendicular to the instrument scale or the observer reads the instrument from an angle. Instruments having a scale and a pointer are normally associated with this type of error. The presence of a mirror behind the pointer or indicator virtually eliminates the occurrence of this type of error.

Effect of misalignment These occur due to the inherent inaccuracies present in the measuring instruments. These errors may also be due to improper use, handling, or selection of the instrument. Wear on the micrometer anvils or anvil faces not being perpendicular to the axis results in misalignment, leading to inaccurate measurements. If the alignment is not proper, sometimes sine and cosine errors also contribute to the inaccuracies of the measurement.

Zero errors When no measurement is being carried out, the reading on the scale of the instrument should be zero. A zero error is defined as that value when the initial value of a physical quantity indicated by the measuring instrument is a non-zero value when it should have actually been zero. For example, a voltmeter might read 1V even when it is not under any electromagnetic influence. This voltmeter indicates 1V more than the true value for all subsequent measurements made. This error is constant for all the values measured using the same instrument. A *constant error* affects all measurements in a measuring process by the same amount or by an amount proportional to the magnitude of the quantity being measured. For

example, in a planimeter, which is used to measure irregular areas, a constant error might occur because of an error in the scale used in the construction of standard or, sometimes, when an incorrect conversion factor is used in conversion between the units embodied by the scale and those in which the results of the measurements are expressed.

Therefore, in order to find out and eliminate any systematic error, it is required to calibrate the measuring instrument before conducting an experiment. Calibration reveals the presence of any systematic error in the measuring instrument.

1.7.2 Random Errors

Random errors provide a measure of random deviations when measurements of a physical quantity are carried out repeatedly. When a series of repeated measurements are made on a component under similar conditions, the values or results of measurements vary. Specific causes for these variations cannot be determined, since these variations are unpredictable and uncontrollable by the experimenter and are random in nature. They are of variable magnitude and may be either positive or negative. When these repeated measurements are plotted, they follow a normal or Gaussian distribution. Random errors can be statistically evaluated, and their mean value and standard deviation can be determined. These errors scatter around a mean value. If n measurements are made using an instrument, denoted by v_1, v_2, v_3, …, v_n, then arithmetic mean is given as

$$\bar{v} = \frac{v_1 + v_2 + v_3 \ldots \ldots \ldots v_n}{n}$$

and standard deviation s is given by the following equation:

$$s = \pm\sqrt{\sum \frac{\left(v - \bar{v}\right)^2}{n}}$$

Standard deviation is a measure of dispersion of a set of readings. It can be determined by taking the root mean square deviation of the readings from their observed numbers, which is given by the following equation:

$$s = \pm\sqrt{\sum \frac{\left(v_1 - \bar{v}\right)^2 + \left(v_2 - \bar{v}\right)^2 + \ldots + \left(v_n - \bar{v}\right)^2}{n}}$$

Random errors can be minimized by calculating the average of a large number of observations. Since precision is closely associated with the repeatability of the measuring process, a precise instrument will have very few random errors and better repeatability. Hence, random errors limit the precision of the instrument. The following are the likely sources of random errors:

1. Presence of transient fluctuations in friction in the measuring instrument
2. Play in the linkages of the measuring instruments
3. Error in operator's judgement in reading the fractional part of engraved scale divisions
4. Operator's inability to note the readings because of fluctuations during measurement
5. Positional errors associated with the measured object and standard, arising due to small variations in setting

Figure 1.5 clearly depicts the relationship between systematic and random errors with respect to the measured value. The measure of a system's accuracy is altered by both systematic and random errors. Table 1.1 gives the differences between systematic and random errors.

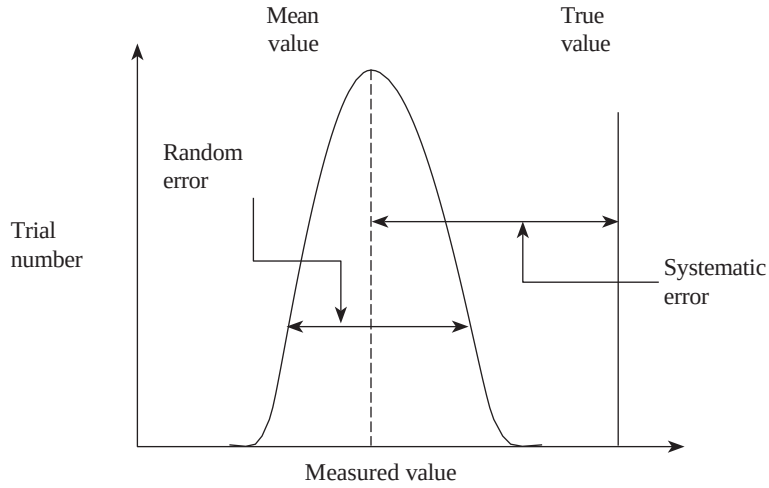


Fig. 1.5 Relationship between systematic and random errors with measured value

Table 1.1 Differences between systematic and random errors

Systematic error	Random error
Not easy to detect	Easy to detect
Cannot be eliminated by repeated measurements	Can be minimized by repeated measurements
Can be assessed easily	Statistical analysis required
Minimization of systematic errors increases the accuracy of measurement	Minimization of random errors increases repeatability and hence precision of the measurement
Calibration helps reduce systematic errors	Calibration has no effect on random errors
Characterization not necessary	Characterized by mean, standard deviation, and variance
Reproducible inaccuracies that are consistently in the same direction	Random in nature and can be both positive and negative

1.8 METHODS OF MEASUREMENT

When precision measurements are made to determine the values of a physical variable, different methods of measurements are employed. Measurements are performed to determine the magnitude of the value and the unit of the quantity under consideration. For instance, the length of a rod is 3 m, where the number, 3, indicates the magnitude and the unit of measurement is metre. The choice of the method of measurement depends on the required accuracy and the amount of permissible error. Irrespective of the method used, the primary objective is to

minimize the uncertainty associated with measurement. The common methods employed for making measurements are as follows:

Direct method In this method, the quantity to be measured is directly compared with the primary or secondary standard. Scales, vernier callipers, micrometers, bevel protractors, etc., are used in the direct method. This method is widely employed in the production field. In the direct method, a very slight difference exists between the actual and the measured values of the quantity. This difference occurs because of the limitation of the human being performing the measurement.

Indirect method In this method, the value of a quantity is obtained by measuring other quantities that are functionally related to the required value. Measurement of the quantity is carried out directly and then the value is determined by using a mathematical relationship. Some examples of indirect measurement are angle measurement using sine bar, measurement of strain induced in a bar due to the applied force, determination of effective diameter of a screw thread, etc.

Fundamental or absolute method In this case, the measurement is based on the measurements of base quantities used to define the quantity. The quantity under consideration is directly measured, and is then linked with the definition of that quantity.

Comparative method In this method, as the name suggests, the quantity to be measured is compared with the known value of the same quantity or any other quantity practically related to it. The quantity is compared with the master gauge and only the deviations from the master gauge are recorded after comparison. The most common examples are comparators, dial indicators, etc.

Transposition method This method involves making the measurement by direct comparison, wherein the quantity to be measured (V) is initially balanced by a known value (X) of the same quantity; next, X is replaced by the quantity to be measured and balanced again by another known value (Y). If the quantity to be measured is equal to both X and Y, then it is equal to

$$V = \sqrt{XY}$$

An example of this method is the determination of mass by balancing methods and known weights.

Coincidence method This is a differential method of measurement wherein a very minute difference between the quantity to be measured and the reference is determined by careful observation of the coincidence of certain lines and signals. Measurements on vernier calliper and micrometer are examples of this method.

Deflection method This method involves the indication of the value of the quantity to be measured directly by deflection of a pointer on a calibrated scale. Pressure measurement is an example of this method.

Complementary method The value of the quantity to be measured is combined with a known value of the same quantity. The combination is so adjusted that the sum of these two values is

equal to the predetermined comparison value. An example of this method is determination of the volume of a solid by liquid displacement.

Null measurement method In this method, the difference between the value of the quantity to be measured and the known value of the same quantity with which comparison is to be made is brought to zero.

Substitution method It is a direct comparison method. This method involves the replacement of the value of the quantity to be measured with a known value of the same quantity, so selected that the effects produced in the indicating device by these two values are the same. The Borda method of determining mass is an example of this method.

Contact method In this method, the surface to be measured is touched by the sensor or measuring tip of the instrument. Care needs to be taken to provide constant contact pressure in order to avoid errors due to excess constant pressure. Examples of this method include measurements using micrometer, vernier calliper, and dial indicator.

Contactless method As the name indicates, there is no direct contact with the surface to be measured. Examples of this method include the use of optical instruments, tool maker's microscope, and profile projector.

Composite method The actual contour of a component to be checked is compared with its maximum and minimum tolerance limits. Cumulative errors of the interconnected elements of the component, which are controlled through a combined tolerance, can be checked by this method. This method is very reliable to ensure interchangeability and is usually effected through the use of composite GO gauges. The use of a GO screw plug gauge to check the thread of a nut is an example of this method.

A QUICK OVERVIEW

- The importance and necessity of metrology have greatly increased with the industrial revolution, which emphasizes the importance of metrology in industries.
- Inspection is defined as a procedure in which a part or product characteristic, for example a dimension, is examined to determine whether it conforms to the design specification. Basically, inspection is carried out to isolate and evaluate a specific design or to measure the quality attribute of a component or product.
- We know that accuracy of measurement is very important in order to manufacture a quality product. *Accuracy* is the degree of agreement of the measured dimension with its true magnitude. It can also be defined as the maximum amount by which the result differs from true value.
- *Precision* is the degree of repetitiveness of the measuring process. It is the degree of agreement of the repeated measurements of a quantity made by using the same method, under similar conditions.
- The ability of the measuring instrument to repeat the same results during the act of measurements for the same quantity is known as *repeatability*.
- The difference between the true value and the mean value of the set of readings on the same component is termed as an *error*. Error can also



Standards of Measurement

After studying this chapter, the reader will be able to

- understand the importance of and need for standards
- describe the evolution of standards and the role of the National Physical Laboratory
- elucidate the different material standards, their contribution, and disadvantages, and appreciate the significance of wavelength standards
- compare the characteristics of line and end standards
- explain the transfer from line standard to end standard
- calibrate end bars

2.1 INTRODUCTION

Human beings have always been innovative and have exploited the natural resources of the earth to manufacture products and devices that satisfy their basic needs and desires. They have always experimented with the form, size, and performance of the products they have invented. During the medieval period, the measurement process underwent an evolution and people accepted the process in specific trades, but no common standards were set. Generally, these measurement standards were region-dependent, and as trade and commerce grew, the need for standardization was also realized. In fact, it would be impossible to imagine today's modern world without a good system of standards of measurement.

Mass production, an idea generated during the last industrial revolution, has become very popular and synonymous with the present manufacturing industry, and a necessity for manufacturing identical parts. Today, almost all manufacturing units practise the principle of interchangeability of manufacture. In order to accomplish complete interchangeability of manufacture in industries, it is essential to a have a measurement system that can adequately define the features of the components/products to the accuracy required.

2.2 STANDARDS AND THEIR ROLES

In order to make measurements a meaningful exercise, some sort of comparison with a known quantity is very essential. It is necessary to define a unit value of any physical quantity under

consideration such that it will be accepted internationally. It is not sufficient to only define these unit values of physical quantities; these should also be measureable. A standard is defined as the fundamental value of any known physical quantity, as established by national and international organizations of authority, which can be reproduced. Fundamental units of physical quantities such as length, mass, time, and temperature form the basis for establishing a measurement system.

In the present world of globalization, it is impossible to perform trade in national and international arenas without standards. In fact, a good system of standards is essential for fair international trade and commerce; it also helps accomplish complete interchangeability of manufacture. Adhering to globally accepted standards allows the manufacturer to convince the customers about the quality of the product.

Standards play a vital role for manufacturers across the world in achieving consistency, accuracy, precision, and repeatability in measurements and in supporting the system that enables the manufacturers to make such measurements.

2.3 EVOLUTION OF STANDARDS

The need for accurate measurements has been recognized by human beings since early times, which is evident from the history of standards. Length standards were one of the earliest standards set by human beings. A brief look at history reveals the following interesting facts. The Egyptian cubit was the earliest recorded length standard—the length of the Pharaoh's forearm plus the width of his palm. The royal cubit was the first master standard made out of black granite, which was used in Egyptian pyramid construction. The actual foot length of the Greek monarch was defined as a foot. King Henry I decreed the length from the tip of nose to the end of the middle finger, when the arm is fully stretched as one yard.

One of the important prerequisites for progress in science is to gain mastery over the science of measurement. Any progress made in manufacturing and other business sectors in the international arena necessitates that activities at high scientific and technical levels be accomplished by progress in metrology. Further, automation in manufacturing industries demands a very high level of accuracy, precision, and reliability. It is pertinent to mention here that human beings' knowledge of nature and the universe, their adaptability to the purpose, and their ability to measure precisely, provides a basis for the science of metrology.

The metric system, which was accepted by France in 1795, coexisted with medieval units until 1840, when it was declared as the exclusive system of weights and measures. In 1798, Eli Whitney introduced the concept of manufacturing interchangeable components for assembling guns. This led to the development of standardization of manufacturing activities to achieve interchangeability. Based on a 4-year investigation, John Quincy Adams in 1821 submitted a report on the metric system and the modernization of our measurement system to the United States Congress. Highlighting the importance of measurement in his report, he stated, 'Weights and measures may be ranked among the necessaries of life to every individual of human society. They enter into the economical arrangements and daily concerns of every family. They are necessary to every occupation of human industry; to the distribution and security of every species of property; to every transaction of trade and commerce; to the labors of the husbandman; to the ingenuity of the artificer; to the studies of the philosopher; to the researches of the antiquarian; to the navigation of the mariner, and the marches of the soldier; to all the

exchanges of peace, and all the operations of war. The knowledge of them, as in established use, is among the first elements of education, and is often learned by those who learn nothing else, not even to read and write. This knowledge is riveted in the memory by the habitual application of it to the employments of men throughout life'.

By 1860, in order to keep pace with scientific inventions, there arose a need for better metric standards. In 1855, the imperial standard yard was developed in England, which was quite accurate. In 1872, the first international prototype metre was developed in France. The International Metric Convention, which was held in France in 1875, universally accepted the metric system, and provisions were also made to set up the International Bureau of Weights and Measures (BIPM) in Paris, which was signed by 17 countries. This convention also agreed to precisely define the metric standards for mass and length and established permanent mechanisms to propose and implement further improvements in the metric system. In 1866, the USA passed an act of Congress to employ the metric system of weights and measures in all contracts, dealings, and court proceedings. In the USA, since 1893, the internationally accepted metric standards have served as the basic measurement standards. Around 35 countries, including continental Europe and most of South America, officially adopted the metric system in 1900.

The internationally adopted standards were required to extend support to the rapid increase in trade between industrialized countries. This resulted in the establishment of international organizations for standardization such as the International Electrotechnical Commission (IEC) in 1906 and the International Organization for Standardization (ISO) in 1947.

In October 1960, at the 11th General Conference on Weights and Measures held in Paris, the original metric standards were redefined in accordance with the 20th-century standards of measurement and a new revised and simplified international system of units, namely the SI units was devised. SI stands for *systeme international d'unites* (international system of units). The seven basic units established under the SI unit system are given in Table 2.1. The 11th General Conference recommended a new standard of length, known as wavelength standard, according to which, metre is defined as 16,50,763.73 × wavelengths of the red–orange radiation of krypton 86 atom in vacuum.

In the 17th General Conference of Weights and Measures held on 20 October 1983, the modern metre was defined as the length of the path travelled by light in vacuum during a time interval of 1/29,97,92,458 of a second.

Table 2.1 Basic units of SI system

Quantity	Unit	Symbol
Length	Metre	m
Mass	Kilogram	kg
Time	Second	s
Thermodynamic temperature	Kelvin	K
Amount of substance	Mole	mol
Electric current	Ampere	A
Luminous intensity	Candela	Cd

In 1999, a mutual recognition arrangement (MRA) was signed by the Committee of Weights and Measures to cater to the growing need for an open, transparent, and comprehensive method to provide users with reliable quantitative information on the comparability of national metrology services. It also provides the technical basis for wider agreements negotiated for international trade, commerce, and regulatory affairs.

In 1999, directors of the National Metrology Institutes (NMIs) signed an MRA for national measurement standards and for calibration and measurement certificates issued by NMIs.

2.4 NATIONAL PHYSICAL LABORATORY

The National Physical Laboratory (NPL) was established in UK in 1900. It is a public institution for standardizing and verifying instruments, testing materials, and determining physical constants. NPL India (NPLI) was established in 1947 in New Delhi under the Council of Scientific and Industrial Research (CSIR). It also has to comply with the statutory obligation of realizing, establishing, maintaining, reproducing, and updating the national standards of measurement and calibration facilities for different parameters.

The main purpose of establishing NPLI is to reinforce and carry out research and development activities in the areas of physical sciences and key physics-based technologies. NPLI is also responsible for maintaining national standards of measurements and ensuring that they conform to international standards. It is established to support industries and national and private agencies in their research and development activities by carrying out calibration and testing, precision measurements, and development of processes and devices. It also ascertains that the national standards of measurements are traceable to the international standards. NPLI also shoulders the responsibility of assisting in research and development activities in the fields of material development, radio and atmospheric sciences, superconductivity and cryogenics, etc.

The major exercise of NPLI is to compare at regular intervals, the national standards with the corresponding standards maintained by the NMIs of other countries in consultation with the International Committee of Weights and Measures and the member nations of the Asia Pacific Metrology Programme. This exercise is essential to establish equivalence of national standards of measurement at NPL with those at other NMIs so that the calibration certificates issued by NPL would have global acceptability.

2.5 MATERIAL STANDARD

Two standard systems for linear measurement that have been accepted and adopted worldwide are English and metric (yard and metre) systems. Most countries have realized the importance and advantages of the metric system and accepted metre as the fundamental unit of linear measurement.

Scientists across the world have always been in the pursuit of a suitable unit for length, and consistent efforts have been made to keep the unit of length constant irrespective of the environmental conditions. The problem with material standards used earlier was that the materials used for defining the standards could change their size with temperature and other conditions. In order to keep the fundamental unit unchanged, great care and attention had to be exercised to maintain the same conditions. The natural and invariable unit for length was finalized as the primary standard when they found that wavelength of monochromatic light was not affected by environmental conditions. They were able to express easily the previously defined yard and metre in terms of the wavelength of light.

Yard or metre is defined as the distance between two scribed lines on a bar of metal maintained under certain conditions of temperature and support. These are legal line standards and are governed by the Act of Parliament for their use.

2.5.1 Yard

The imperial standard yard is a bronze bar 1 sq. inch in cross-section and 38 inches in length, having a composition of 82% Cu, 13% tin, and 5% Zn. The bar contains holes of ½-inch diameter × ½-inch depth. It has two round recesses, each located one inch away from either end and extends up to the central plane of the bar. A highly polished gold plug having a diameter of 1/10 of an inch comprises three transversely engraved lines and two longitudinal lines that are inserted into each of these holes such that the lines lie in the neutral plane. The top surface of the plug lies on the neutral axis. Yard is then defined as the distance between the two central transverse lines of the plug maintained at a temperature of 62 °F. Yard, which was legalized in 1853, remained a legal standard until it was replaced by the wavelength standard in 1960. One of the advantages of maintaining the gold plug lines at neutral axis is that this axis remains unaffected due to bending of the beam. Another advantage is that the gold plug is protected from getting accidentally damaged. Three orthographic views of the imperial standard yard are shown in Fig. 2.1. It is important to note that an error occurs in the neutral axis because of the support provided at the ends. This error can be minimized by placing the supports in such a way that the slope at the ends is zero and the flat end faces of the bar are mutually parallel to each other.

Airy points can be defined as the points at which a horizontal rod is optionally supported to prevent it from bending. These points are used to support a length standard in such a way as to minimize the error due to bending. Sir George Biddell Airy (1801–92) showed that the distance d between the supports can be determined by the formula $\frac{1}{\sqrt{n^2-1}} \times L$, where n is the number of supports and L, the length of the bar. When it is supported by two points, $n = 2$. Substituting this value in the preceding formula, the distance between the supports is obtained as $0.577L$. This means that the supports should be at an equal distance from each end to a position where they are $0.577L$ apart. Normally, airy points are marked for length bars greater than 150 mm.

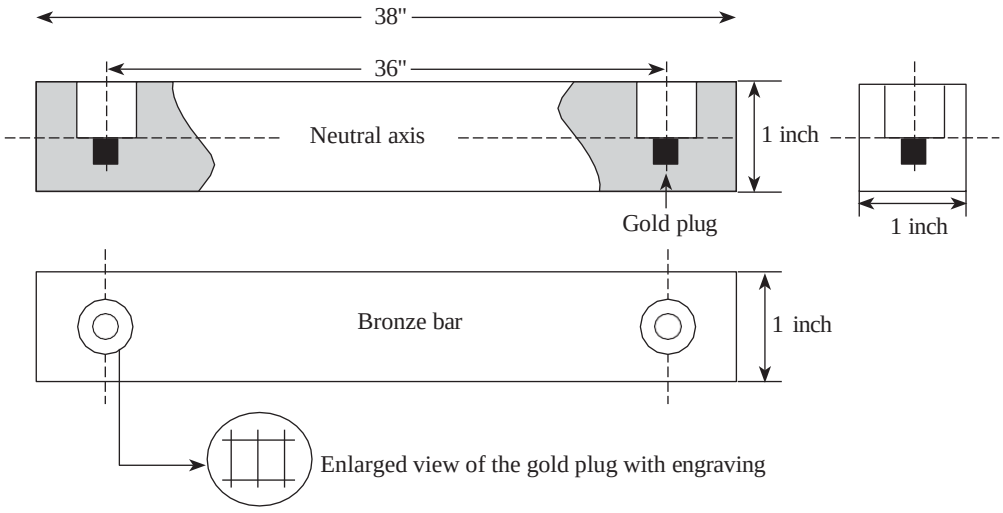


Fig. 2.1 Imperial standard yard

The distance between two supports for international yard and international prototype metre is marked as 29.94 inches and 58.9 mm, respectively.

2.5.2 Metre

This standard is also known as international prototype metre, which was established in 1875. It is defined as the distance between the centre positions of the two lines engraved on the highly polished surface of a 102 cm bar of pure platinum–iridium alloy (90% platinum and 10% iridium) maintained at 0 °C under normal atmospheric pressure and having the cross-section of a web, as shown in Fig. 2.2. The top surface of the web contains graduations coinciding with the neutral axis of the section. The web-shaped section offers two major advantages. Since the section is uniform and has graduations on the neutral axis, it allows the whole surface to be graduated. This type of cross-section provides greater rigidity for the amount of metal involved and is economical even though an expensive metal is used for its construction. The bar is inoxidizable and can have a good polish, which is required for obtaining good-quality lines. It is supported by two rollers having at least 1 cm diameter, which are symmetrically located in the same horizontal plane at a distance of 751 mm from each other such that there is minimum deflection.

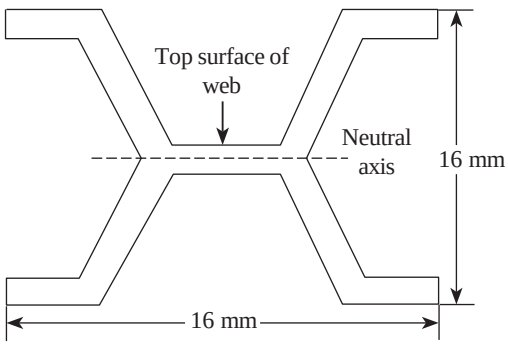


Fig. 2.2 International prototype metre

2.5.3 Disadvantages of Material Standards

The following disadvantages are associated with material standards:

1. Material standards are affected by changes in environmental conditions such as temperature, pressure, humidity, and ageing, resulting in variations in length.
2. Preservation of these standards is difficult because they must have appropriate security to prevent their damage or destruction.
3. Replicas of material standards are not available for use at other places.
4. They cannot be easily reproduced.
5. Comparison and verification of the sizes of gauges pose considerable difficulty.
6. While changing to the metric system, a conversion factor is necessary.

These disadvantages prompted scientists to look for a standard that remains unaffected by the variations in environmental conditions. Consequently, the 11th General Conference of Weights and Measures, which was held in Paris in 1960, recommended a new standard of length, known as wavelength standard, measured in terms of wavelengths of the red–orange radiation of the krypton 86 isotope gas.

2.6 WAVELENGTH STANDARD

It is very clear from the methods discussed earlier that comparison and verification of the sizes of the gauges pose considerable difficulty. This difficulty arises because the working

standard used as a reference is derived from a physical standard and successive comparisons are required to establish the size of a working standard using the process discussed earlier, leading to errors that are unacceptable. By using wavelengths of a monochromatic light as a natural and invariable unit of length, the dependency of the working standard on the physical standard can be eliminated. The definition of a standard of length relative to the metre can easily be expressed in terms of the wavelengths of light.

The use of the interference phenomenon of light waves to provide a working standard may thus be accepted as ultimate for all practical purposes. However, there were some objections to the use of the light wavelength standard because of the impossibility of producing pure monochromatic light, as wavelength depends upon the amount of isotope impurity in the elements. However, with rapid advancements in the field of atomic energy, pure isotopes of natural elements have been produced. Cadmium 114, krypton 86, and mercury 198 are possible sources of radiation of wavelengths suitable for the natural standard of length. There is no need to preserve the wavelength standard as it is not a physical one. This standard of length is reproducible, and the error of reproduction can be of the order of 1 part in 100 million.

Finally, in 1960, at the 11th General Conference of Weights and Measures held in Paris, it was recommended and decided that krypton 86 is the most suitable element if used in a hot-cathode discharge lamp maintained at a temperature of 68 K. According to this standard, metre is defined as 1,650,763.73 × wavelengths of the red–orange radiation of a krypton 86 atom in vacuum. This standard can be reproduced with an accuracy of about 1 part in 10^9 and can be accessible to any laboratory.

2.6.1 Modern Metre

The modern metre was defined in the 17th General Conference of Weights and Measures held on 20 October 1983. According to this, the metre is the length of the path travelled by light in vacuum during a time interval of 1/299,792,458 of a second. This standard is technologically more accurate and feasible when compared to the red–orange radiation of a krypton 86 atom and can be realized in practice through the use of an iodine-stabilized helium–neon laser. The reproducibility of the modern metre is found to be 3 parts in 10^{11}, which could be compared to measuring the earth's mean circumference to an accuracy of about 1 mm.

2.7 SUBDIVISIONS OF STANDARDS

The imperial standard yard and metre defined in Sections 2.5.1 and 2.5.2 are master standards that cannot be used for daily measurement purposes. In order to facilitate measurement at different locations depending upon the relative importance of standard, they are subdivided into the following four groups:

Primary standards For defining the unit precisely, there shall be one and only one material standard. Primary standards are preserved carefully and maintained under standard atmospheric conditions so that they do not change their values. This has no direct application to a measuring problem encountered in engineering. These are used only for comparing with secondary standards. International yard and international metre are examples of standard units of length.

Secondary standards These are derived from primary standards and resemble them very closely with respect to design, material, and length. Any error existing in these bars is recorded by comparison with primary standards after long intervals. These are kept at different locations under strict supervision and are used for comparison with tertiary standards (only when it is absolutely essential). These safeguard against the loss or destruction of primary standards.

Tertiary standards Primary and secondary standards are the ultimate controls for standards; these are used only for reference purposes and that too at rare intervals. Tertiary standards are reference standards employed by NPL and are used as the first standards for reference in laboratories and workshops. These standards are replicas of secondary standards and are usually used as references for working standards.

Working standards These are used more frequently in workshops and laboratories. When compared to the other three standards, the materials used to make these standards are of a lower grade and cost. These are derived from fundamental standards and suffer from loss of instrumental accuracy due to subsequent comparison at each level in the hierarchical chain. Working standards include both line and end standards.

Accuracy is one of the most important factors to be maintained and should always be traceable to a single source, usually the national standards of the country. National laboratories of most of the developed countries are in close contact with the BIPM. This is essential because ultimately all these measurements are compared with the standards developed and maintained by the bureaus of standards throughout the world. Hence, there is an assurance that items manufactured to identical dimensions in different countries will be compatible with each other, which helps in maintaining a healthy trade. Figure 2.3 shows the hierarchical classification of standards and Table 2.2 shows a classification based on purpose. The accuracy of a particular standard depends on subsequent comparisons of standards in the hierarchical ladder. While assessing the accuracy of a standard it is also important to consider factors such as the care taken during comparison, the

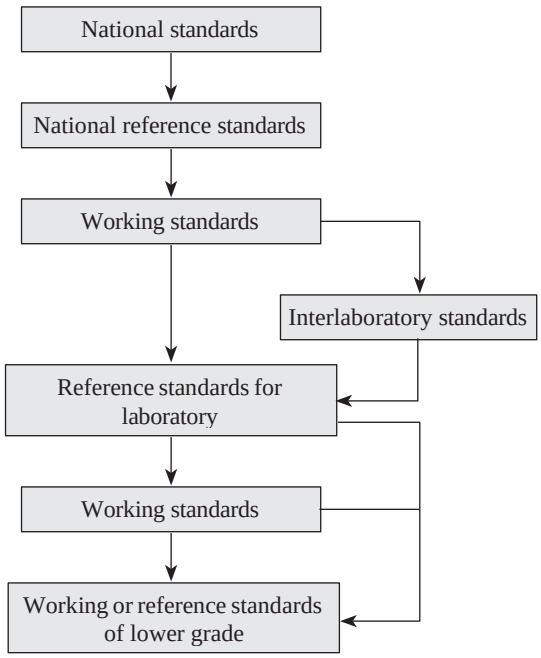


Fig. 2.3 Hierarchical classification of standards

Table 2.2 Classification of standards based on purpose

Standard	Purpose
Reference	Reference
Calibration	Calibration of inspection and working standards
Inspection	Used by inspectors
Working standards	Used by operators

procedure followed for comparison when the recent comparison was made, and the stability of the particular standard. It is imperative that there is a degradation of accuracy in the standards that follow the national standards in the hierarchical classification of standards, as shown in Fig. 2.3.

Comparisons with national standards are seldom performed, as frequent comparisons may degrade their accuracy. For frequent comparisons of the lower-order standards, national reference standards are generally used. For calibration purposes, normally working standards are employed. These working standards stand third in the hierarchical classification and are sometimes called national working standards. Interlaboratory standards and reference standards for laboratories, which are of good quality, are derived from working standards.

2.8 LINE AND END MEASUREMENTS

We all know that sometimes distances have to be measured between two lines or two surfaces or between a line and a surface. When the distance between two engraved lines is used to measure the length, it is called line standard or line measurement. The most common examples are yard and metre. The rule with divisions marked with lines is widely used.

When the distance between two flat parallel surfaces is considered a measure of length, it is known as end standard or end measurement. The end faces of the end standards are hardened to reduce wear and lapped flat and parallel to a very high degree of accuracy. The end standards are extensively used for precision measurement in workshops and laboratories. The most common examples are measurements using slip gauges, end bars, ends of micrometer anvils, vernier callipers, etc.

For an accurate measurement, it is required to select a measuring device that suits a particular measuring situation. For example, for a direct measurement of the distances between two edges, a rule is not suitable because it is a line-measuring device. However, a comparison of characteristics of the line and end standards clearly shows that the end standards provide higher accuracy than line standards.

2.8.1 Characteristics of Line Standards

The following are the characteristics of line standards:

1. Measurements carried out using a scale are quick and easy and can be used over a wide range.
2. Even though scales can be engraved accurately, it is not possible to take full advantage of this accuracy. The engraved lines themselves possess thickness, making it difficult to perform measurements with high accuracy.
3. The markings on the scale are not subjected to wear. *Undersizing* occurs as the leading ends are subjected to wear.
4. A scale does not have a built-in datum, which makes the alignment of the scale with the axis of measurement difficult. This leads to undersizing.
5. Scales are subjected to parallax effect, thereby contributing to both positive and negative reading errors.
6. A magnifying lens or microscope is required for close tolerance length measurement.

2.8.2 Characteristics of End Standards

End standards comprise a set of standard blocks or bars using which the required length is created. The characteristics of these standards are as follows:

1. These standards are highly accurate and ideal for making close tolerance measurement.
2. They measure only one dimension at a time, thereby consuming more time.
3. The measuring faces of end standards are subjected to wear.
4. They possess a built-in datum because their measuring faces are flat and parallel and can be positively located on a datum surface.
5. Groups of blocks/slip gauges are wrung together to create the required size; faulty wringing leads to inaccurate results.
6. End standards are not subjected to parallax errors, as their use depends on the feel of the operator.
7. Dimensional tolerance as close as 0.0005 mm can be obtained.

The end and line standards are initially calibrated at 20 ± ½ ºC. Temperature changes influence the accuracy of these standards. Care should be taken in the manufacturing of end and line standards to ensure that change of shape with time is minimum or negligible. Table 2.3 gives a complete comparison of line and end standards.

Table 2.3 Comparison of line and end standards

Characteristics	Line standard	End standard
Principle of measurement	Distance between two engraved lines is used as a measure of length	Distance between two flat and parallel surfaces is used as a measure of length
Accuracy of measurement	Limited accuracy of ±0.2 mm; magnifying lens or microscope is required for high accuracy	High accuracy of measurement; close tolerances upto ±0.0005 mm can be obtained
Ease and time of measurement	Measurements made using a scale are quick and easy	Measurements made depend on the skill of the operator and are time consuming
Wear	Markings on the scale are not subjected to wear. Wear may occur on leading ends, which results in undersizing	Measuring surfaces are subjected to wear
Alignment	Alignment with the axis of measurement is not easy, as they do not contain a built-in datum	Alignment with the axis of measurement is easy, as they possess a built-in datum
Manufacture	Manufacturing process is simple	Manufacturing process is complex
Cost	Cost is low	Cost is high
Parallax effect	Subjected to parallax effect	No parallax error; their use depends on the feel of the operator
Wringing	Does not exist	Slip gauges are wrung together to build the required size
Examples	Scale (yard and metre)	Slip gauges, end bars, ends of micrometer anvils, and vernier callipers

2.8.3 Transfer from Line Standard to End Standard

We know that primary standards are basically line standards and that end standards are practical workshop standards. Line standards are highly inconvenient for general measurement purposes and are usually used to calibrate end standards, provided that the length of primary line standard is accurately known. There is a probability of the existence of a very small error in the primary standard, which may not be of serious concern. It is important to accurately determine the error in the primary standard so that the lengths of the other line standards can be precisely evaluated when they are compared with it.

From the aforementioned discussions it is clear that when measurements are made using end standards, the distance is measured between the working faces of the measuring instrument, which are flat and mutually parallel. A composite line standard is used to transfer a line standard to an end standard.

Figure 2.4 shows a primary line standard of a basic length of 1 m whose length is accurately known. A line standard having a basic length of more than 1 m is shown in Fig. 2.5. This line standard consists of a central length bar that has a basic length of 950 mm. Two end blocks of 50 mm each are wrung on either end of the central bar. Each end block contains an engraved line at the centre.

The composite line standard whose length is to be determined is compared with the primary line standard, and length L is obtained as using the following formula:

$$L = L_1 + b + c$$

The four different ways in which the two end blocks can be arranged using all possible combinations and then compared with the primary line standard are

$$L = L_1 + b + c$$
$$L = L_1 + b + d$$
$$L = L_1 + a + c$$
$$L = L_1 + a + d$$

Summation of these four measurements gives

$$4L = 4L_1 + 2a + 2b + 2c + 2d$$
$$= 4L_1 + 2(a + b) + 2(c + d)$$

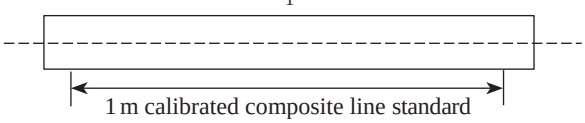


Fig. 2.4 Basic length of primary line standard

Now, the combination of blocks $(a + b)$ and $(c + d)$ are unlikely to be of the same length. The two are therefore compared; let the difference between them be x, as shown in Fig. 2.6.

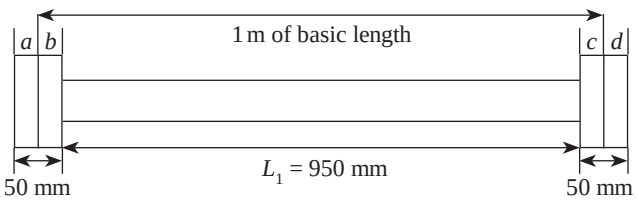


Fig. 2.5 Basic length of primary line standard with end blocks

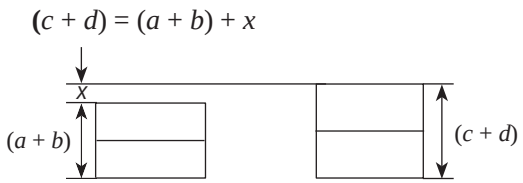


Fig. 2.6 Comparison of blocks $(a + b)$ and $(c + d)$

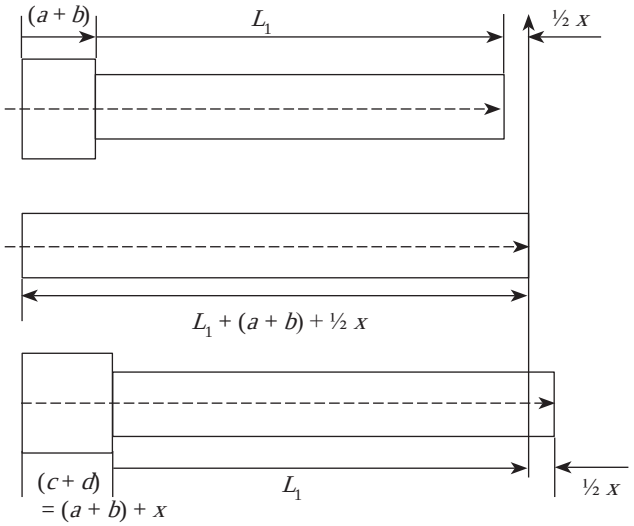


Fig. 2.7 End standard of known length as obtained from the standard

Substituting the value of $(c + d)$,

$$4L = 4L_1 + 2(a + b) + 2[(a + b) + x)]$$
$$4L = 4L_1 + 2(a + b) + 2(a + b) + 2x$$
$$4L = 4L_1 + 4(a + b) + 2x$$

Dividing by 4, we get

$$L = L_1 + (a + b) + \tfrac{1}{2}x$$

An end standard of known length can now be obtained consisting of either $L_1 + (a + b)$ or $L_1 + (c + d)$, as shown in Fig. 2.7. The length of $L_1 + (a + b)$ is $L_1 + (a + b) + \frac{1}{2}x$ less $\frac{1}{2}x$, where $(a + b)$ is shorter of the two end blocks. The length of $L_1 + (c + d)$ is $L_1 + (a + b) + \frac{1}{2}x$ plus $\frac{1}{2}x$, where $(c + d)$ is longer of the two end blocks. The calibrated composite end bar can be used to calibrate a solid end standard of the same basic length.

2.9 BROOKES LEVEL COMPARATOR

The Brookes level comparator (Fig. 2.8) is used to calibrate standards by comparing with a master standard. End standards can be manufactured very accurately using a Brookes level comparator. A.J.C. Brookes devised this simple method in 1920 and hence the name. The Brookes level comparator has a very accurate spirit level. In order to achieve an accurate comparison, the spirit level is supported on balls so that it makes only a point contact with the gauges.

The table on which the gauges are placed for comparison are first levelled properly using the spirit level. The two gauges (the master standard gauge and the standard gauge) that are

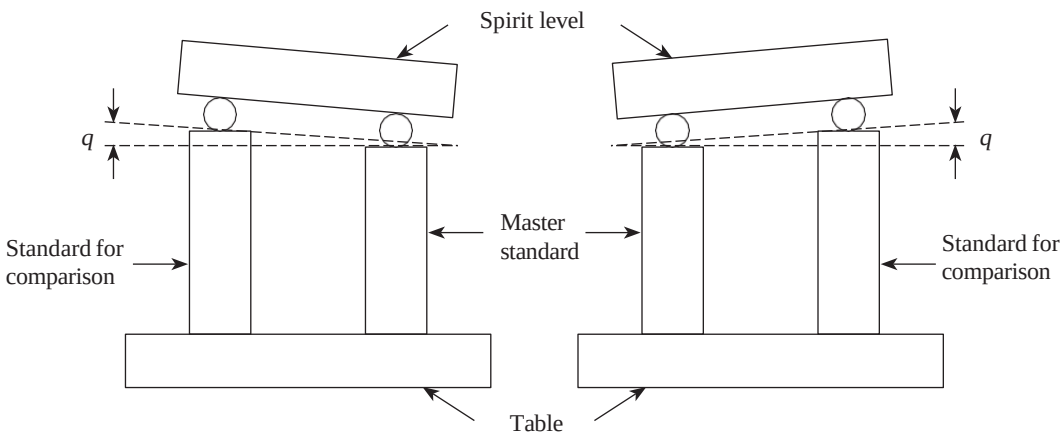


Fig. 2.8 Calibration using a Brookes level comparator

to be compared are wrung on the table and the spirit level is properly placed on them. The bubble reading is recorded at this position. The positions of the two gauges are interchanged by rotating the table by 180°. The spirit level is again placed to note down the bubble reading at this position. The arrangement is shown in Fig. 2.8.

The two readings will be the same if the gauges are of equal length and different for gauges of unequal lengths. When the positions of the gauges are interchanged, the level is tilted through an angle equal to twice the difference in the height of gauges divided by the spacing of level supports. The bubble readings can be calibrated in terms of the height difference, as the distance between the two balls is fixed. The effect of the table not being levelled initially can be eliminated because of the advantage of turning the table by 180°.

2.10 DISPLACEMENT METHOD

The displacement method is used to compare an edge gauge with a line standard. This method is schematically represented in Fig. 2.9. The line standard, which is placed on a carrier, is positioned such that line A is under the cross-wires of a fixed microscope, as seen in Fig. 2.9(a). The spindle of the micrometer is rotated until it comes in contact with the projection on the carrier and then the micrometer reading is recorded. The carrier is moved again to position line B under the cross-wires of the microscope. At this stage, the end gauge is inserted as shown in Fig. 2.9(b) and the micrometer reading is recorded again. Then the sum of the length of the line standard and the difference between the micrometer readings will be equal to the length of the end gauge.

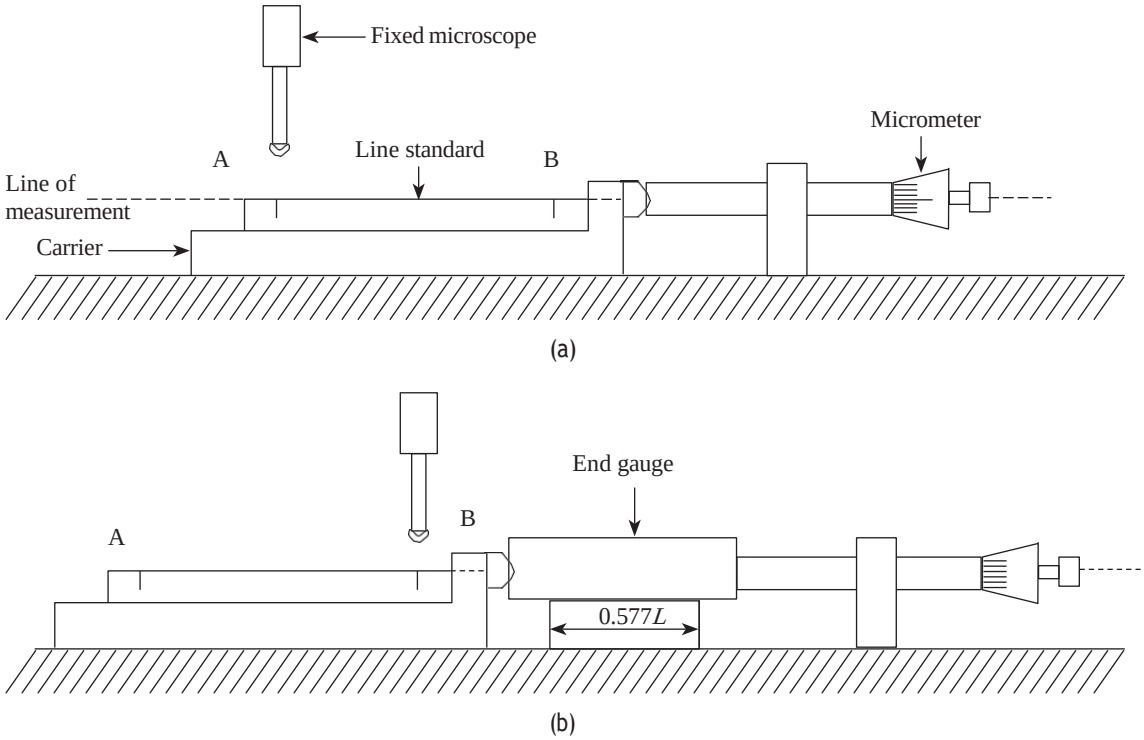


Fig. 2.9 Displacement method (a) With line standard only (b) With line and end standards combined

2.11 CALIBRATION OF END BARS

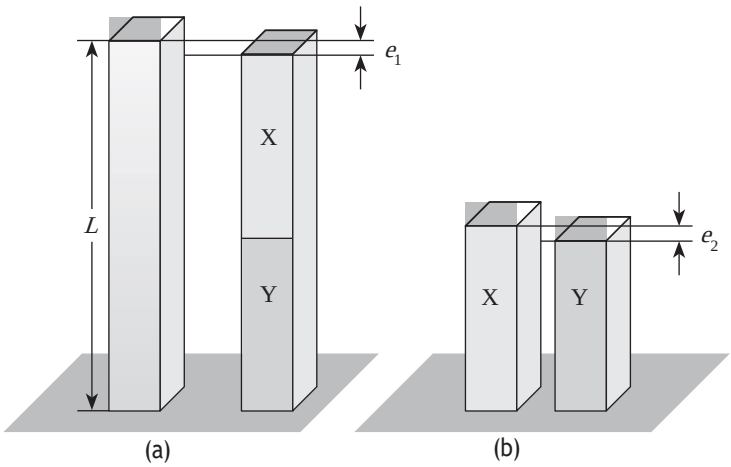


Fig. 2.10 Calibration of end bars (a) Comparison of metre bar and end bars wrung together (b) Comparison of individual end bars

In order to calibrate two bars having a basic length of 500 mm with the help of a one piece metre bar, the following procedure is adopted.

The metre bar to be calibrated is wrung to a surface plate. The two 500 mm bars to be calibrated are wrung together to form a bar that has a basic length of 1 m, which in turn is wrung to the surface plate beside the metre bar, as shown in Fig. 2.10(a). The difference in height e_1 is obtained. The two 500 mm bars are then compared to determine the difference in length, as shown in Fig. 2.10(b).

Let L_X and L_Y be the lengths of the two 500 mm bars. Let e_1 be the difference in height between the calibrated metre bar and the combined lengths of X and Y. Let the difference between the lengths of X and Y be e_2. Let L be the actual length of the metre bar.

Then the first measurement gives a length of $L \pm e_1 = L_X + L_Y$, depending on whether the combined length of L_X and L_Y is longer or shorter than L.

The second measurement yields a length of $L_X \pm e_2 = L_Y$, again depending on whether X is longer or shorter than Y.

Then substituting the value of L_Y from the second measurement in the first measurement, we get

$$L \pm e_1 = L_X + L_X \pm e_2 = 2L_X \pm e_2$$

or

$$2L_X = L \pm e_1 \pm e_2$$

Therefore, $L_X = (L \pm e_1 \pm e_2)/2$ and $L_Y = L_X \pm e_2$

For calibrating three, four, or any other number of length standards of the same basic size, the same procedure can be followed. One of the bars is used as a reference while comparing the individual bars and the difference in length of the other bar is obtained relative to this bar.

2.12 NUMERICAL EXAMPLES

Example 2.1 It is required to obtain a metre standard from a calibrated line standard using a composite line standard. The actual length of the calibrated line standard is 1000.015 mm. The composite line standard comprises a length bar having a basic length of 950 mm and two end blocks, (a + b) and (c + d), each having a basic length of 50 mm. Each end block contains an engraved line at the centre.

Four different measurements were obtained when comparisons were made between the calibrated line standard and the composite bar using all combinations of end blocks: L_1 = 1000.0035 mm, L_2 = 1000.0030 mm, L_3 = 1000.0020 mm, and L_4 = 1000.0015 mm. Determine the actual length of the metre bar.
Block ($a + b$) was found to be 0.001 mm greater than block ($c + d$) when two end blocks were compared with each other.

Solution

The sum of all the four measurements for different combinations is as follows:

$$4L = L_1 + L_2 + L_3 + L_4 = 4L_1 + 4(a + b) + 2x$$

However, $4L$ = 1000.0035 + 1000.0030 + 1000.0020 + 1000.0015 = 4000.0100 mm
Therefore, $4L = 4000.0100 = 4L_1 + 4(a + b) + 2x$
Dividing by 4, we get $1000.0025 = L_1 + (a + b) + ½x$
When block ($a + b$) is used, the end standard length $= L_1 + (a + b) + ½x$
$= 1000.0025 + 0.0005 = 1000.0030$ mm
When block ($c + d$) is used, the end standard length $= L_1 + (a + b) - ½x$
$= 1000.0025 - 0.0005 = 1000.0020$ mm

Example 2.2 A calibrated metre end bar, which has an actual length of 1000.0005 mm, is to be used in the calibration of two bars X and Y, each having a basic length of 500 mm. When compared with the metre bar, the sum of L_X and L_Y is found to be shorter by 0.0003 mm. When X and Y are compared, it is observed that X is 0.0004 mm longer than Y. Determine the actual length of X and Y.

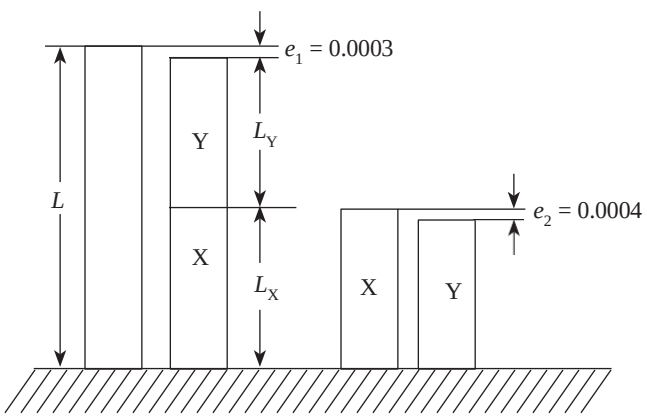


Fig. 2.11 Arrangement of end bars

Solution

The arrangement of gauges is as shown in Fig. 2.11.
From first principles, we have

$$L - e_1 = L_X + L_Y$$

However, $L_X = L_Y + e_2$
Therefore, we get $L - e_1 = 2L_Y + e_2$
Or, $L_Y = (L - e_1 - e_2)/2$
Substituting the values for L, e_1, and e_2, we get

$$L_Y = (1000.0005 - 0.0003 - 0.0004)/2$$
$$L_Y = 999.9998/2 \text{ mm}$$

i.e., L_Y = 499.9999 mm

We have $L_X = L_Y + e_2$

$$L_X = 499.9999 + 0.0004 = 500.0003 \text{ mm}$$

Example 2.3 Three 200 mm gauges to be calibrated are measured on a level comparator by wringing them together and then comparing them with a 600 mm gauge. The 600 mm gauge has an actual length of 600.0025 mm, and the three gauges together have a combined length of 600.0035 mm. When the three gauges are intercompared, it is found that gauge A is longer than gauge B by 0.0020 mm but shorter than gauge C by 0.001 mm. Determine the length of each gauge.

Copyrighted Materials
Copyright © 2013 Oxford University Press Retrieved from www.knovel.com

Limits, Fits, and Tolerances

After studying this chapter, the reader will be able to

- understand the importance of manufacturing components to specified sizes
- elucidate the different approaches of interchangeable and selective assembly
- appreciate the significance of different types of limits, fits, and tolerances in design and manufacturing fields, which are required for efficient and effective performance of components/products
- utilize the principle of limit gauging and its importance in inspection in industries
- design simple GO and NOT gauges used in workshops/ inspection

3.1 INTRODUCTION

Although any two things found in nature are seldom identical, they may be quite similar. This is also true in the case of manufacturing of different components for engineering applications. No two parts can be produced with identical measurements by any manufacturing process. A manufacturing process essentially comprises five m's—man, machine, materials, money, and management.

Variations in any of the first three elements induce a change in the manufacturing process. All the three elements are subjected to natural and characteristic variations. In any production process, regardless of how well it is designed or how carefully it is maintained, a certain amount of natural variability will always exist. These natural variations are random in nature and are the cumulative effect of many small, essentially uncontrollable causes. When these natural variations in a process are relatively small, we usually consider this to be an acceptable level of process performance.

Usually, variability arises from improperly adjusted machines, operator error, tool wear, and/or defective raw materials. Such characteristic variability is generally large when compared to the natural variability. This variability, which is not a part of random or chance cause pattern, is referred to as 'assignable causes'. Characteristic variations can be attributed to assignable causes that can easily be identified and controlled. However, this has to be

achieved economically, which brings in the fourth element. Characteristic variability causes variations in the size of components. If the process can be kept under control, that is, all the assignable and controllable causes of variations have been eliminated or controlled, the size variations will be well within the prescribed limits. These variations can be modified through operator or management action.

Production processes must perform consistently to meet the production and design requirements. In order to achieve this, it is essential to keep the process under control. Thus, when the process is under control, distribution of most of the measured values will be around the mean value in a more or less symmetrical way, when plotted on a chart. It is therefore impossible to produce a part to an exact size or basic size and some variations, known as tolerances, need to be allowed. Some variability in dimension within certain limits must be tolerated during manufacture, however precise the process may be. The permissible level of tolerance depends on the functional requirements, which cannot be compromised.

No component can be manufactured precisely to a given dimension; it can only be made to lie between two limits, upper (maximum) and lower (minimum). The designer has to suggest these tolerance limits, which are acceptable for each of the dimensions used to define shape and form, and ensure satisfactory operation in service. When the tolerance allowed is sufficiently greater than the process variation, no difficulty arises. The difference between the upper and lower limits is termed *permissive tolerance.*

For example, a shaft has to be manufactured to a diameter of 40 ± 0.02 mm. This means that the shaft, which has a basic size of 40 mm, will be acceptable if its diameter lies anywhere between the limits of sizes, that is, an upper limit of 40.02 mm and a lower limit of 39.98 mm. Then permissive tolerance is equal to $40.02 - 39.98 = 0.04$. *Basic* or *nominal size* is defined as the size based on which the dimensional deviations are given.

In any industry, a manufactured product consists of many components. These components when assembled should have a proper fit, in order for the product to function properly and have an extended life. Fit depends on the correct size relationships between the two mating parts. Consider the example of rotation of a shaft in a hole. Enough clearance must be provided between the shaft and the hole to allow an oil film to be maintained for lubrication purpose. If the clearance is too small, excessive force would be required to rotate the shaft. On the other hand, if the clearance is too wide, there would be vibrations and rapid wear resulting in ultimate failure. Therefore, the desired clearance to meet the requirements has to be provided. Similarly, to hold the shaft tightly in the hole, there must be enough interference between the two so that forces of elastic compression grip them tightly and do not allow any relative movement between them.

An ideal condition would be to specify a definite size to the hole and vary the shaft size for a proper fit or vice versa. Unfortunately, in practice, particularly in mass production, it is not possible to manufacture a part to the exact size due to the inherent inaccuracy of manufacturing methods. Even if a part is manufactured to the exact size by chance, it is not possible to measure it accurately and economically during machining. In addition, attempts to manufacture to the exact size can increase the production cost.

Dimensional variations, although extremely small, do exist because of the inevitable inaccuracies in tooling, machining, raw material, and operators. If efforts are made to identify and reduce or eliminate common causes of variation, that is, if the process is kept under control, then the resultant frequency distribution of dimensions produced will have a normal

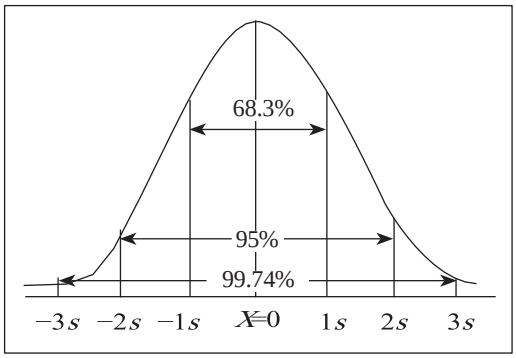


Fig. 3.1 Normal or Gaussian frequency distribution

or Gaussian distribution, that is, 99.74% parts will be well within ±3*s* limits of the mean value, as shown in Fig. 3.1. Thus, it is possible to express the uncertainty in measurement as a multiple of standard deviation. This value is determined by the probability that a measurement will fall outside the stated limits.

3.2 PRINCIPLE OF INTERCHANGEABILITY

For manufacturing a large number of components, it is not economical to produce both the mating parts (components) using the same operator. Further, such parts need to be manufactured within minimum possible time without compromising on quality. To enable the manufacture of identical parts, mass production, an idea of the last industrial revolution that has become very popular and synonymous with the present manufacturing industry, becomes inevitable.

Modern production techniques require that a complete product be broken into various component parts so that the production of each part becomes an independent process, leading to specialization. The various components are manufactured in one or more batches by different persons on different machines at different locations and are then assembled at one place.

To achieve this, it is essential that the parts are manufactured in bulk to the desired accuracy and, at the same time, adhere to the limits of accuracy specified. Manufacture of components under such conditions is called *interchangeable manufacture*.

When interchangeable manufacture is adopted, any one component selected at random should assemble with any other arbitrarily chosen mating component. In order to assemble with a predetermined fit, the dimensions of the components must be confined within the permissible tolerance limits. By *interchangeable assembly*, we mean that identical components, manufactured by different operators, using different machine tools and under different environmental conditions, can be assembled and replaced without any further modification during the assembly stage and without affecting the functioning of the component when assembled. Production on an interchangeable basis results in an increased productivity with a corresponding reduction in manufacturing cost. Modern manufacturing techniques that complement mass production of identical parts facilitating interchangeability of components have been developed. When components are produced in bulk, unless they are interchangeable, the purpose of mass production is not fulfilled.

For example, consider the assembly of a shaft and a part with a hole. The two mating parts are produced in bulk, say 1000 each. By interchangeable assembly any shaft chosen randomly should assemble with any part with a hole selected at random, providing the desired fit.

Another major advantage of interchangeability is the ease with which replacement of defective or worn-out parts is carried out, resulting in reduced maintenance cost. In addition, the operator, by performing the same limited number of operations, becomes a specialist in that work. By achieving specialization in labour, there will be a considerable reduction in manufacturing and assembly time and enhancement in quality. Interchangeable manufacture increases productivity and reduces production and time costs.

tolerances desired in such industries are very narrow and impossible to achieve economically by any sophisticated machine tools. Balls are segregated into different groups depending on their size to enable the assembly of any bearing with balls of uniform size. In a broader sense, a combination of both interchangeable and selective assemblies exists in modern-day manufacturing industries, which help to manufacture quality products.

3.3 TOLERANCES

To satisfy the ever-increasing demand for accuracy, the parts have to be produced with less dimensional variation. Hence, the labour and machinery required to manufacture a part has become more expensive. It is essential for the manufacturer to have an in-depth knowledge of the tolerances to manufacture parts economically but, at the same time, adhere to quality and reliability aspects. In fact, precision is engineered selectively in a product depending on the functional requirements and its application. To achieve an increased compatibility between mating parts to enable interchangeable assembly, the manufacturer needs to practise good tolerance principles. Therefore, it is necessary to discuss some important principles of tolerances that are usually employed for manufacturing products.

We know that it is not possible to precisely manufacture components to a given dimension because of the inherent inaccuracies of the manufacturing processes. The components are manufactured in accordance with the permissive tolerance limits, as suggested by the designer, to facilitate interchangeable manufacture. The permissible limits of variations in dimensions have to be specified by the designer in a logical manner, giving due consideration to the functional requirements. The choice of the tolerances is also governed by other factors such as manufacturing process, cost, and standardization.

Tolerance can be defined as the magnitude of permissible variation of a dimension or other measured value or control criterion from the specified value. It can also be defined as the total variation permitted in the size of a dimension, and is the algebraic difference between the upper and lower acceptable dimensions. It is an absolute value.

The basic purpose of providing tolerances is to permit dimensional variations in the manufacture of components, adhering to the performance criterion as established by the specification and design. If high performance is the sole criterion, then functional requirements dictate the specification of tolerance limits; otherwise, the choice of setting tolerance, to a limited extent, may be influenced and determined by factors such as methods of tooling and available manufacturing equipment. The industry follows certain approved accuracy standards, such as ANSI (American National Standards Institute) and ASME (American Society of Mechanical Engineers), to manufacture different parts.

3.3.1 Computer-aided Modelling

Nowadays, computers are widely being employed in the design and manufacture of parts. Most leading design tools such as AEROCADD, AUTOCAD, and Solid Works, which are currently being used in industries, are equipped with tolerance features. The algorithms and programming codes that are in existence today are aimed at enhancing the accuracy with minimum material wastage. These programs have the capability of allotting tolerance ranges for different miniature parts of complex mechanical systems.



Fig. 3.2 Relationship between work tolerance and manufacturing cost

3.3.2 Manufacturing Cost and Work Tolerance

It is very pertinent to relate the production of components within the specified tolerance zone to its associated manufacturing cost. As the permissive tolerance goes on decreasing, the manufacturing cost incurred to achieve it goes on increasing exponentially. When the permissive tolerance limits are relaxed without degrading the functional requirements, the manufacturing cost decreases. This is clearly illustrated in Fig. 3.2.

Further, in order to maintain such close tolerance limits, manufacturing capabilities have to be enhanced, which certainly increases the manufacturing cost. The components manufactured have to undergo a closer scrutiny, which demands stringent inspection procedures and adequate instrumentation. This increases the cost of inspection. Hence, tolerance is a trade-off between the economical production and the accuracy required for proper functioning of the product. In fact, the tolerance limits specified for the components to be manufactured should be just sufficient to perform their intended functions.

3.3.3 Classification of Tolerance

Tolerance can be classified under the following categories:

1. Unilateral tolerance
2. Bilateral tolerance
3. Compound tolerance
4. Geometric tolerance

Unilateral Tolerance

When the tolerance distribution is only on one side of the basic size, it is known as unilateral tolerance. In other words, tolerance limits lie wholly on one side of the basic size, either above or below it. This is illustrated in Fig. 3.3(a). Unilateral tolerance is employed when precision fits are required during assembly. This type of tolerance is usually indicated when the mating parts are also machined by the same operator. In this system, the total tolerance as related to the basic size is in one direction only. Unilateral tolerance is employed in the drilling process wherein dimensions of the hole are most likely to deviate in one direction only, that is, the hole is always oversized rather than undersized. This system is preferred because the basic size is used for the GO limit gauge. This helps in standardization of the GO gauge, as holes and shafts of different grades will have the same lower and upper limits, respectively. Changes in the magnitude of the tolerance affect only the size of the other gauge dimension, the NOT GO gauge size.

Example

$$40^{+0.02}_{+0.01},\ 40^{+0.02}_{-0.00},\ 40^{-0.01}_{-0.02},\ 40^{+0.00}_{-0.02}$$

Bilateral Tolerance

When the tolerance distribution lies on either side of the basic size, it is known as bilateral tolerance. In other words, the dimension of the part is allowed to vary on both sides of the basic

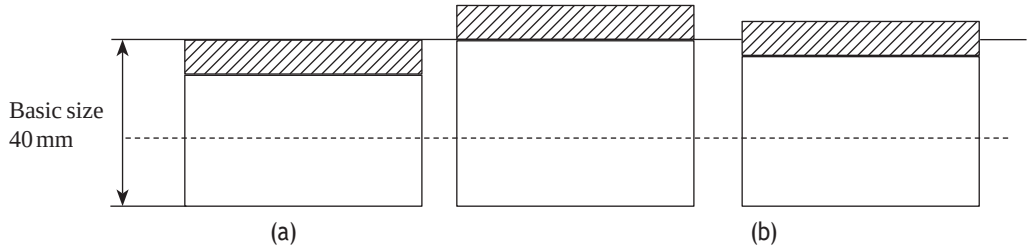


Fig. 3.3 Tolerances (a) Unilateral (b) Bilateral

size but may not be necessarily equally disposed about it. The operator can take full advantage of the limit system, especially in positioning a hole. This system is generally preferred in mass production where the machine is set for the basic size. This is depicted in Fig. 3.3(b). In case unilateral tolerance is specified in mass production, the basic size should be modified to suit bilateral tolerance.

Example

$$40 \pm 0.02,\ 40\ ^{+0.02}_{-0.01}$$

Compound Tolerance

When tolerance is determined by established tolerances on more than one dimension, it is known as compound tolerance For example, tolerance for the dimension *R* is determined by the combined effects of tolerance on 40 mm dimension, on 60º, and on 20 mm dimension. The tolerance obtained for dimension *R* is known as compound tolerance (Fig. 3.4). In practice, compound tolerance should be avoided as far as possible.

Geometric Tolerance

Normally, tolerances are specified to indicate the actual size or dimension of a feature such as a hole or a shaft. In order to manufacture components more accurately or with minimum dimensional variations, the manufacturing facilities and the labour required become more cost intensive. Hence, it is essential for the manufacturer to have an in-depth knowledge of tolerances, to manufacture quality and reliable components economically. In fact, depending on the application of the end product, precision is engineered selectively. Therefore, apart from considering the actual size, other geometric dimensions such as roundness and straightness of a shaft have to be considered while manufacturing components. The tolerances specified should also encompass such variations. However, it is difficult to combine all errors of roundness, straightness, and diameter within a single tolerance on diameter. Geometric tolerance is defined as the total amount that the dimension of a manufactured part can vary. Geometric tolerance underlines the importance of the shape of a feature as against its size. Geometric dimensioning and tolerancing is a method of defining parts based on how they function, using standard symbols. This method is frequently used in industries. Depending on the functional requirements, tolerance on

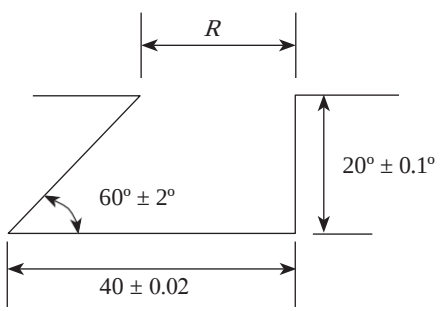


Fig. 3.4 Compound tolerance

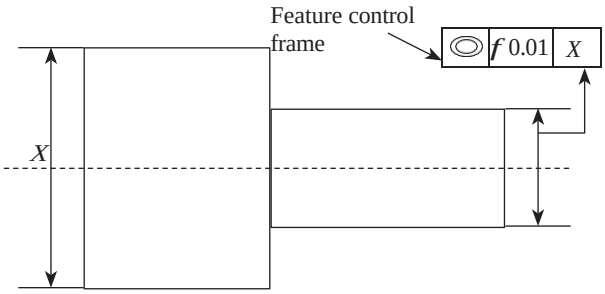


Fig. 3.5 Representation of geometric tolerance

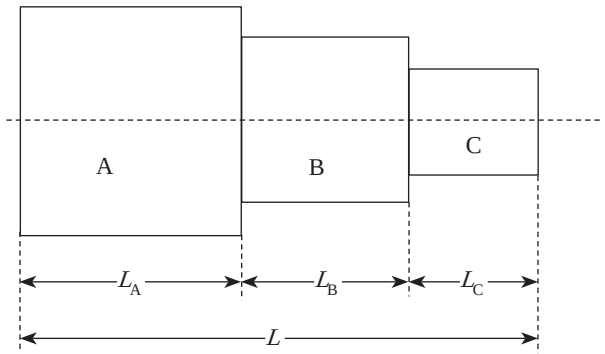


Fig. 3.6 Accumulation of tolerances

diameter, straightness, and roundness may be specified separately. Geometric tolerance can be classified as follows:

Form tolerances Form tolerances are a group of geometric tolerances applied to individual features. They limit the amount of error in the shape of a feature and are independent tolerances. Form tolerances as such do not require locating dimensions. These include straightness, circularity, flatness, and cylindricity.

Orientation tolerances Orientation tolerances are a type of geometric tolerances used to limit the direction or orientation of a feature in relation to other features. These are related tolerances. Perpendicularity, parallelism, and angularity fall into this category.

Positional tolerances Positional tolerances are a group of geometric tolerances that controls the extent of deviation of the location of a feature from its true position. This is a three-dimensional geometric tolerance comprising position, symmetry, and concentricity.

Geometric tolerances are used to indicate the relationship of one part of an object with another. Consider the example shown in Fig. 3.5. Both the smaller- and the larger-diameter cylinders need be concentric with each other. In order to obtain a proper fit between the two cylinders, both the centres have to be in line with each other. Further, perhaps both the cylinders are manufactured at different locations and need to be assembled on an interchangeable basis. It becomes imperative to indicate how much distance can be tolerated between the centres of these two cylinders. This information can be represented in the feature control frame that comprises three boxes. The first box on the left indicates the feature to be controlled, which is represented symbolically. In this example, it is concentricity. The box at the centre indicates the distance between the two cylinders that can be tolerated, that is, these two centres cannot be apart by more than 0.01 mm. The third box indicates that the datum is with X. The different types of geometric tolerances and their symbolic representation are given in Table 3.1.

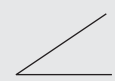
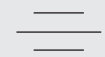
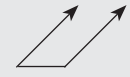
Consider the example shown in Fig. 3.6.

Let $L_A = 30^{+0.02}_{-0.01}$ mm, $L_B = 20^{+0.02}_{-0.01}$ mm, and $L_C = 10^{+0.02}_{-0.01}$ mm.

The overall length of the assembly is the sum of the individual length of components given as

$$L = L_A + L_B + L_C$$
$$L = 30 + 20 + 10 = 60\,\text{mm}$$

Table 3.1 Symbolic representation of geometric tolerances

Type of geometric tolerance	Feature	Geometric characteristic	Definition	Symbol
Form tolerance	Independent	Straightness (two-dimensional)	Controls the extent of deviation of a feature from a straight line	
		Circularity (two-dimensional)	Exercises control on the extent of deviation of a feature from a perfect circle	
		Flatness (three-dimensional)	Controls the extent of deviation of a feature from a flat plane	
		Cylindricity (three-dimensional)	Controls the extent of deviation of a feature from a perfect cylinder	
	Related or single	Profile of a line (two-dimensional)	Controls the extent of deviation of an outline of a feature from the true profile	
		Profile of a surface (two-dimensional)	Exercises control on the extent of deviation of a surface from the true profile	
Orientation tolerance	Related	Perpendicularity (three-dimensional)	Exercises control on the extent of deviation of a surface, axis, or plane from a 90° angle	
		Parallelism (three-dimensional)	Controls the extent of deviation of a surface, axis, or plane from an orientation parallel to the specified datum	
		Angularity (three-dimensional)	Exercises controls on the deviation of a surface, axis, or plane from the angle described in the design specifications	
Position tolerance	Related	Position (three-dimensional)	Exercises control on the extent of deviation of location of a feature from its true position	
		Symmetry (three-dimensional)	Exercises control on extent of deviation of the median points between two features from a specified axis or centre plane	
		Concentricity (three-dimensional)	Exercises control on extent of deviation of the median points of multiple diameters from the specified datum axis	
Runout	Related	Circular runout (two-dimensional)	Exercises composite control on the form, orientation, and location of multiple cross-sections of a cylindrical part as it rotates	
		Total runout (three-dimensional)	Exercises simultaneous composite control on the form, orientation, and location of the entire length of a cylindrical part as it rotates	

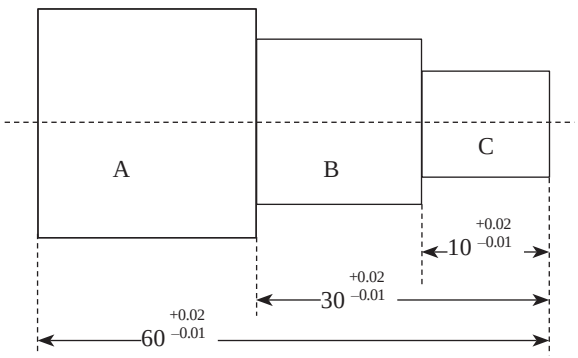


Fig. 3.7 Progressive dimensioning

Then, cumulative upper tolerance limit is 0.02 + 0.02 + 0.02 = 0.06 mm and cumulative lower limit = − 0.01 − 0.01 − 0.01 = −0.03 mm

Therefore, dimension of the assembled length will be = $60^{+0.06}_{-0.03}$ mm

It is essential to avoid or minimize the cumulative effect of tolerance build-up, as it leads to a high tolerance on overall length, which is undesirable. If progressive dimensioning from a common reference line or a baseline dimensioning is adopted, then tolerance accumulation effect can be minimized. This is clearly illustrated in Fig. 3.7.

3.4 MAXIMUM AND MINIMUM METAL CONDITIONS

Let us consider a shaft having a dimension of 40 ± 0.05 mm.

The maximum metal limit (MML) of the shaft will have a dimension of 40.05 mm because at this higher limit, the shaft will have the maximum possible amount of metal.

The shaft will have the least possible amount of metal at a lower limit of 39.95 mm, and this limit of the shaft is known as minimum or least metal limit (LML).

Similarly, consider a hole having a dimension of 45 ± 0.05 mm.

The hole will have a maximum possible amount of metal at a lower limit of 44.95 mm and the lower limit of the hole is designated as MML. For example, when a hole is drilled in a component, minimum amount of material is removed at the lower limit size of the hole. This lower limit of the hole is known as MML.

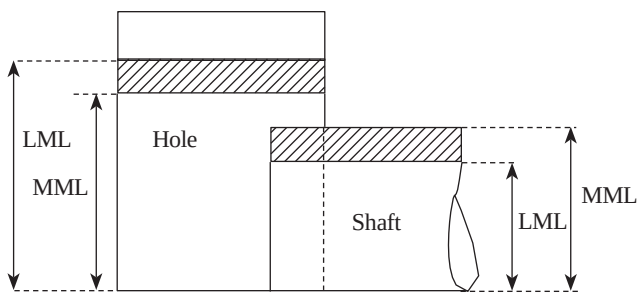


Fig. 3.8 MML and LML

The higher limit of the hole will be the LML. At a high limit of 45.05 mm, the hole will have the least possible amount of metal. The maximum and minimum metal conditions are shown in Fig. 3.8.

3.5 FITS

Manufactured parts are required to mate with one another during assembly. The relationship between the two mating parts that are to be assembled, that is, the hole and the shaft, with respect to the difference in their dimensions before assembly is called a *fit.* An ideal fit is required for proper functioning of the mating parts. Three basic types of fits can be identified, depending on the actual limits of the hole or shaft:

1. Clearance fit

2. Interference fit
3. Transition fit

Clearance fit The largest permissible diameter of the shaft is smaller than the diameter of the smallest hole. This type of fit always provides clearance. Small clearances are provided for a precise fit that can easily be assembled without the assistance of tools. When relative motions are required, large clearances can be provided, for example, a shaft rotating in a bush. In case of clearance fit, the difference between the sizes is always positive. The clearance fit is described in Fig. 3.9.

Interference fit The minimum permissible diameter of the shaft exceeds the maximum allowable diameter of the hole. This type of fit always provides interference. Interference fit is a form of a tight fit. Tools are required for the precise assembly of two parts with an interference fit. When two mating parts are assembled with an interference fit, it will be an almost permanent assembly, that is, the parts will not come apart or move during use. To assemble the parts with interference, heating or cooling may be required. In an interference fit, the difference between the sizes is always negative. Interference fits are used when accurate location is of utmost importance and also where such location relative to another part is critical, for example, alignment of dowel pins. The interference fit is illustrated in Fig. 3.10.

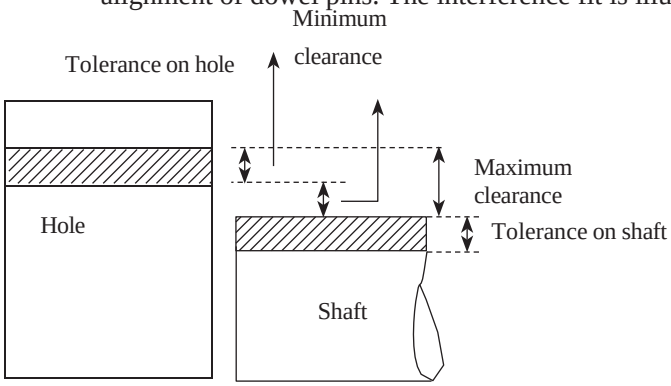


Fig. 3.9 Clearance fit

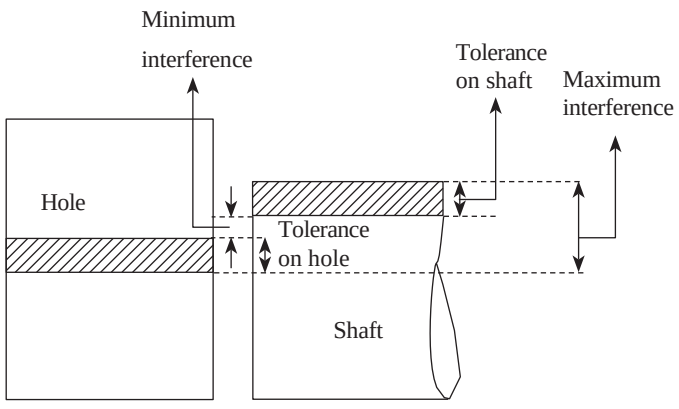


Fig. 3.10 Interference fit

Transition fit The diameter of the largest permissible hole is greater than the diameter of the smallest shaft and the diameter of the smallest hole is smaller than the diameter of the largest shaft. In other words, the combination of maximum diameter of the shaft and minimum diameter of the hole results in an interference fit, while that of minimum diameter of the shaft and maximum diameter of the hole yields a clearance fit. Since the tolerance zones overlap, this type of fit may sometimes provide clearance and sometimes interference, as depicted in Fig. 3.11. Precise assembly may be obtained with the assistance of tools, for example, dowel pins may be required in tooling to locate parts.

In a clearance fit, minimum clearance is the difference between minimum size of the hole, that is, low limit of the hole (LLH), and maximum size of the shaft, that is, high limit of the shaft (HLS), before assembly. In a transition or a

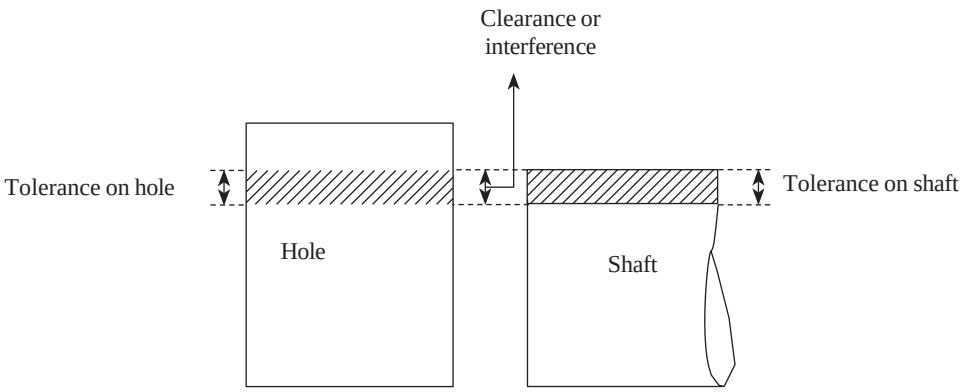


Fig. 3.11 Transition fit

clearance fit, maximum clearance is the arithmetical difference between the maximum size of the hole, that is, high limit of the hole (HLH), and the minimum size of the shaft, that is, low limit of the shaft (LLS), before assembly.

In an interference fit, minimum interference is the arithmetical difference between maximum size of the hole, that is, HLH, and minimum size of the shaft, that is, LLS, before assembly. In a transition or an interference fit, it is the arithmetical difference between minimum size of the hole, that is, LLH, and maximum size of the shaft, that is, HLS, before assembly.

Thus, in order to find out the type of fit, one needs to determine HLH − LLS and LLH − HLS. If both the differences are positive, the fit obtained is a clearance fit, and if negative, it is an interference fit. If one difference is positive and the other is negative, then it is a transition fit.

The three basic types of fits, clearance, transition, and interference, can be further classified, as shown in Fig. 3.12.

3.5.1 Allowance

An allowance is the intentional difference between the maximum material limits, that is, LLH and HLS (minimum clearance or maximum interference) of the two mating parts. It is the

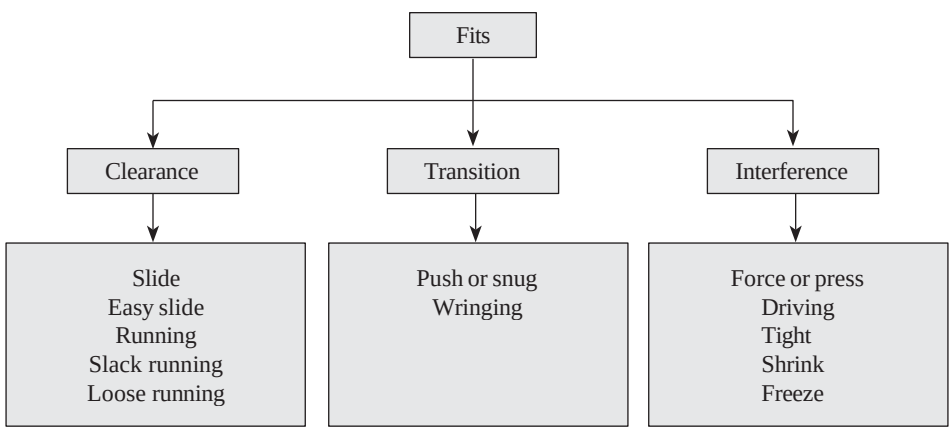


Fig. 3.12 Detailed classification of fits

Table 3.2 Examples of different types of fits

Description of fit	Class of fit	Application area
	Clearance fit	
Slide	H7/h6	Sealing rings, bearing covers, movable gears in change gear trains, clutches, etc.
Easy slide	H7/g7	Lathe spindle, spigots, piston, and slide valves
Running	H8/f8	Lubricated bearings (with oil or grease), pumps and smaller motors, gear boxes, shaft pulleys, etc.
Slack running	H8/c11	Oil seals with metal housings, multi-spline shafts, etc.
Loose running	H8/d9	Loose pulleys, loose bearings with low revolution, etc.
	Interference fit	
Force or press	H8/r6	Crankpins, car wheel axles, bearing bushes in castings, etc.
Driving	H7/s6	Plug or shaft slightly larger than the hole
Tight	H7/p6	Stepped pulleys on the drive shaft of a conveyor
Shrink	H7/u6, H8/u7	Bronze crowns on worm wheel hubs, couplings, gear wheels, and assembly of piston pin in IC engine piston
Freeze	H7/u6, H8/u7	Insertion of exhaust valve seat inserts in engine cylinder blocks and insertion of brass bushes in various assemblies
	Transition fit	
Push or snug	H7/k6	Pulleys and inner ring of ball bearings on shafts
Wringing	H7/n6	Gears of machine tools

prescribed difference between the dimensions of the mating parts to obtain the desired type of fit. Allowance may be positive or negative. Positive allowance indicates a clearance fit, and an interference fit is indicated by a negative allowance.

$$\text{Allowance} = \text{LLH} - \text{HLS}$$

Table 3.2 gives examples of the classification of fits.

3.5.2 Hole Basis and Shaft Basis Systems

To obtain the desired class of fits, either the size of the hole or the size of the shaft must vary. Two types of systems are used to represent the three basic types of fits, namely clearance, interference, and transition fits. They are (a) hole basis system and (b) shaft basis system.

Although both systems are the same, hole basis system is generally preferred in view of the functional properties.

Hole Basis System

In this system, the size of the hole is kept constant and the shaft size is varied to give various types of fits. In a hole basis system, the fundamental deviation or lower deviation of the hole is zero, that is, the lower limit of the hole is the same as the basic size. The two limits of the shaft and the higher dimension of the hole are then varied to obtain the desired type of fit, as illustrated in Fig. 3.13.

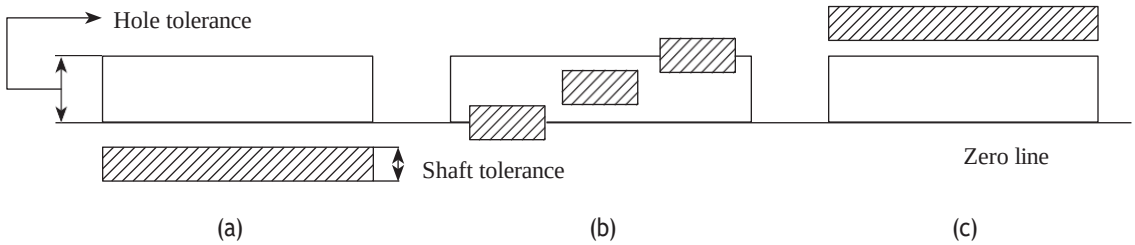


Fig. 3.13 Hole basis system (a) Clearance fit (b) Transition fit (c) Interference fit

This type of system is widely adopted in industries, as it is easier to manufacture shafts of varying sizes to the required tolerances. Standard size drills or reamers can be used to obtain a variety of fits by varying only the shaft limits, which leads to greater economy of production. The shaft can be accurately produced to the required size by standard manufacturing processes, and standard-size plug gauges are used to check hole sizes accurately and conveniently.

Shaft Basis System

The system in which the dimension of the shaft is kept constant and the hole size is varied to obtain various types of fits is referred to as shaft basis system. In this system, the fundamental deviation or the upper deviation of the shaft is zero, that is, the HLH equals the basic size. The desired class of fits is obtained by varying the lower limit of the shaft and both limits of the hole, as shown in Fig. 3.14.

This system is not preferred in industries, as it requires more number of standard-size tools such as reamers, broaches, and gauges, which increases manufacturing and inspection costs. It is normally preferred where the hole dimension is dependent on the shaft dimension and is used in situations where the standard shaft determines the dimensions of the mating parts such as couplings, bearings, collars, gears, and bushings.

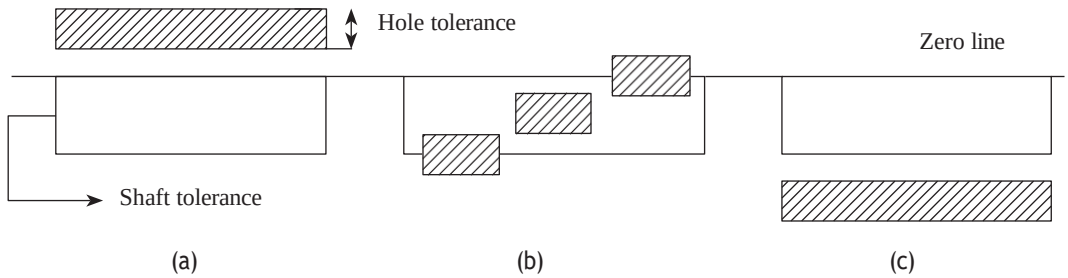


Fig. 3.14 Shaft basis system (a) Clearance fit (b) Transition fit (c) Interference fit

3.5.3 Numerical Examples

Example 3.1 In a limit system, the following limits are specified for a hole and shaft assembly:

$$\text{Hole} = 30^{+0.02}_{+0.00}\ \text{mm and shaft} = 30^{-0.02}_{-0.05}\ \text{mm}$$

Determine the (a) tolerance and (b) allowance.

Solution

(a) Determination of tolerance:

Tolerance on hole = HLH − LLH
= 30.02 − 30.00 = 0.02 mm

Tolerance on shaft = HLS − LLS
= [(30 − 0.02) − (30 − 0.05)] = 0.03 mm

(b) Determination of allowance:

Allowance = Maximum metal condition of hole − Maximum metal condition of shaft
= LLH − HLS
= 30.02 − 29.98 = 0.04 mm

Example 3.2 The following limits are specified in a limit system, to give a clearance fit between a hole and a shaft:

$$\text{Hole} = 25^{+0.03}_{-0.00}\ \text{mm and shaft} = 25^{-0.006}_{-0.020}\ \text{mm}$$

Determine the following:
(a) Basic size
(b) Tolerances on shaft and hole
(c) Maximum and minimum clearances

Solution

(a) Basic size is the same for both shaft and hole.

(b) Determination of tolerance:

Tolerance on hole = HLH − LLH
= 25.03 − 25.00 = 0.03 mm

Tolerance on shaft = HLS − LLS
= [(25 − 0.006) − (25 − 0.020)] = 0.014 mm

Determination of clearances:

Maximum clearance = HLH − LLS
= 25.03 − 24.98 = 0.05 mm

Minimum clearance = LLH − HLS
= 25.00 − (25 − 0.006) = 0.06 mm

Example 3.3 Tolerances for a hole and shaft assembly having a nominal size of 50 mm are as follows:

$$\text{Hole} = 50^{+0.02}_{+0.00}\ \text{mm and shaft} = 50^{-0.05}_{-0.08}\ \text{mm}$$

Determine the following:
(a) Maximum and minimum clearances
(b) Tolerances on shaft and hole
(c) Allowance
(d) MML of hole and shaft
(e) Type of fit

Solution

(a) Determination of clearances:

Maximum clearance = HLH − LLS
= 50.02 − (50 − 0.08) = 0.10 mm

Minimum clearance = LLH − HLS
= 50.00 − (50 − 0.005) = 0.05 mm

(b) Determination of tolerance:

Tolerance on hole = HLH − LLH
= 50.02 − 50.00 = 0.02 mm

Tolerance on shaft = HLS − LLS
= [(50 − 0.05) − (50 − 0.08)] = 0.03 mm

(c) Determination of allowance:

Allowance = Maximum metal condition of hole − Maximum metal condition of shaft
= LLH − HLS
= 50.00 − (50 − 0.05) = 0.05 mm

(d) Determination of MMLs:

MML of hole = Lower limit of hole = 50.00 mm

MML of shaft = Higher limit of shaft = 50.00 − 0.05 = 49.05 mm

(e) Since both maximum and minimum clearances are positive, it can be conclude that the given pair has a clearance fit.

Example 3.4 A clearance fit has to be provided for a shaft and bearing assembly having a diameter of 40 mm. Tolerances on hole and shaft are 0.006 and 0.004 mm, respectively. The tolerances are disposed unilaterally. If an allowance of 0.002 mm is provided, find the limits of size for hole and shaft when (a) hole basis system and (b) shaft basis system are used.

Solution

(a) When hole basis system is used:

Hole size:

HLH = 40.006 mm

LLH = 40.000 mm

The allowance provided is +0.002 mm.

Therefore, HLS = LLH − Allowance
= 40.000 − 0.002 = 39.998 mm

LLS = HLS − Tolerance
= 39.998 − 0.004 = 39.994 mm

(b) When shaft basis system is used:

Shaft size:

HLS = 40.000 mm

LLS = 40.000 − 0.004 = 39.996 mm

The allowance provided is +0.002 mm.

Therefore, LLH = HLS + allowance
= 40.000 + 0.002 = 40.002 mm

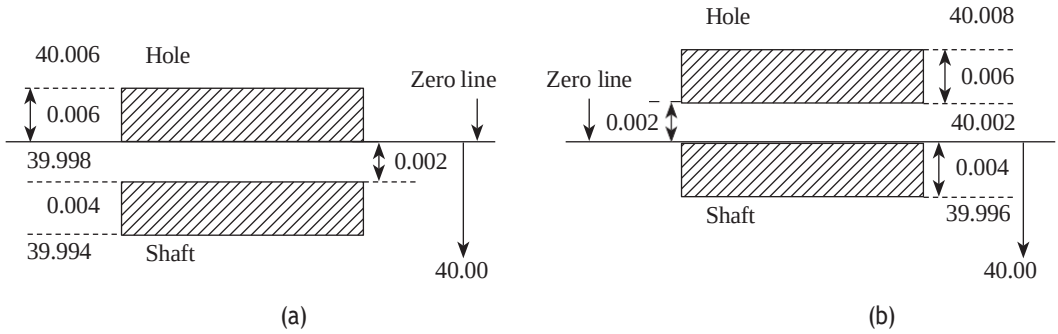


Fig. 3.15 Disposition of tolerances (a) Hole basis system (b) Shaft basis system

HLH = 40.002 + 0.006 = 40.008 mm

The disposition of tolerance for both hole basis and shaft basis systems are given in Fig. 3.15.

Example 3.5 For the following hole and shaft assembly, determine (a) hole and shaft tolerance and (b) type of fit.

$$\text{Hole} = 20^{+0.025}_{+0.000}\ \text{mm and shaft} = 20^{+0.080}_{+0.005}\ \text{mm}$$

Solution

(a) Determination of tolerance:

Tolerance on hole = HLH − LLH
= 20.025 − 20.00 = 0.025 mm

Tolerance on shaft = HLS − LLS
= 20.080 − 20.005 = 0.075 mm

(b) To determine the type of fit, calculate maximum and minimum clearances:

Maximum clearance = HLH − LLS
= 20.025 − 20.005 = 0.020 mm

(Clearance because the difference is positive)

Minimum clearance = LLH − HLS
= 20.00 − 20.080 = −0.08 mm

(Interference because the difference is negative)

Since one difference is positive and the other negative, it can be concluded that the given hole and shaft pair has a transition fit.

Example 3.6 For the following hole and shaft assembly, determine (a) hole and shaft tolerance and (b) type of fit

$$\text{Hole} = 20^{+0.05}_{+0.00}\ \text{mm and shaft} = 20^{+0.08}_{+0.06}\ \text{mm}$$

Solution

(a) Determination of tolerance:

Tolerance on hole = HLH − LLH
= 20.05 − 20.00 = 0.05 mm
Tolerance on shaft = HLS − LLS
= 20.08 − 20.06 = 0.02 mm

(b) To determine the type of fit, calculate maximum and minimum clearances:
Maximum clearance = HLH − LLS
= 20.05 − 20.06 = −0.01 mm
Minimum clearance = LLH − HLS
= 20.00 − 20.08 = −0.08 mm

Since both differences are negative, it can be concluded that the given hole and shaft pair has an interference fit.

3.6 SYSTEM OF LIMITS AND FITS

The rapid growth of national and international trade necessitates the developments of formal systems of limits and fits, at the national and international levels. Economic success of most manufacturing industries critically depends on the conformity of the specifications of the products to international standards. The International Organization for Standardization (ISO) specifies the internationally accepted system of limits and fits. Indian standards are in accordance with the ISO.

The ISO system of limits and fits comprises 18 grades of fundamental tolerances to indicate the level of accuracy of the manufacture. These fundamental tolerances are designated by the letters IT followed by a number. The ISO system provides tolerance grades from IT01, IT0, and IT1 to IT16 to realize the required accuracy. The greater the number, the higher the tolerance limit. The choice of tolerance is guided by the functional requirements of the product and economy of manufacture. The degree of accuracy attained depends on the type and condition of the machine tool used. Table 3.3 gives the fundamental tolerance values required for various applications.

Tolerance values corresponding to grades IT5–IT16 are determined using the standard tolerance unit (i, in μm), which is a function of basic size.

Table 3.3 Tolerances grades for different applications

Fundamental tolerance	Applications
IT01-IT4	For production of gauges, plug gauges, and measuring instruments
IT5-IT7	For fits in precision engineering applications such as ball bearings, grinding, fine boring, high-quality turning, and broaching
IT8-IT11	For general engineering, namely turning, boring, milling, planning, rolling, extrusion, drilling, and precision tube drawing
IT12-IT14	For sheet metal working or press working
IT15-IT16	For processes such as casting, stamping, rubber moulding, general cutting work, and flame cutting

$$i = 0.453\sqrt[3]{D} + 0.001D \text{ microns}$$

where D is the diameter of the part in mm. The linear factor $0.001D$ counteracts the effect of measuring inaccuracies that increase by increasing the measuring diameter. By using this formula, the value of tolerance unit 'i' is obtained for sizes up to 500 mm. D is the geometric mean of the lower and upper diameters of a particular diameter step within which the given or chosen diameter D lies and is calculated by using the following equation:

$$\sqrt{D_{\max} \times D_{\min}}$$

The various steps specified for the diameter steps are as follows:

1–3, 3–6, 6–10, 10–18, 18–30, 30–50, 50–80, 80–120, 120–180, 180–250, 250–315, 315–400, 400–500 500–630, 630–800, and 800–1000 mm.

Tolerances have a parabolic relationship with the size of the products. The tolerance within which a part can be manufactured also increases as the size increases. The standard tolerances corresponding to IT01, IT0, and IT1 are calculated using the following formulae:

IT01: $0.3 + 0.008D$

IT0: $0.5 + 0.012D$

IT1: $0.8 + 0.020D$

The values of tolerance grades IT2–IT4, which are placed between the tolerance grades of IT1 and IT5, follow a geometric progression to allow for the expansion and deformation affecting both the gauges and the workpieces as dimensions increase. For the tolerance grades IT6–IT16, each grade increases by about 60% from the previous one, as indicated in Table 3.4.

Table 3.4 Standard tolerance units

Tolerance grade	IT6	IT7	IT8	IT9	IT10	IT11	IT12	IT13	IT14	IT15	IT16
Standard tolerance unit (*i*)	10	16	25	40	64	100	160	250	400	640	1000

The tolerance zone is governed by two limits: the size of the component and its position related to the basic size. The position of the tolerance zone, from the zero line (basic size), is determined by fundamental deviation. The ISO system defines 28 classes of basic deviations for holes and shafts, which are marked by capital letters A, B, C, ..., ZC (with the exception of I, L, O, Q, and W) and small letters a, b, c, ..., zc (with the exception of i, l, o, q, and w), respectively, as depicted in Fig. 3.16. Different combinations of fundamental deviations and fundamental tolerances are used to obtain various types of fits.

The values of these tolerance grades or fundamental deviations depend on the basic size of the assembly. The different values of standard tolerances and fundamental deviations can be obtained by referring to the design handbook. The choice of the tolerance grade is governed by the type of manufacturing process and the cost associated with it. From Fig. 3.16, a typical case can be observed in which the fundamental deviation for both hole H and shaft h having a unilateral tolerance of a specified IT grade is zero. The first eight designations from A (a) to H (h) for holes (shafts) are intended to be used in clearance fit, whereas the remaining designations, JS (js) to ZC (zc) for holes (shafts), are used in interference or transition fits. For JS, the two deviations are equal and given by ±IT/2.

Consider the designation 40 H7/d9. In this example, the basic size of the hole and shaft is 40 mm. The nature of fit for the hole basis system is designated by H and the fundamental deviation of the hole is zero. The tolerance grade is indicated by IT7. The shaft has a d-type fit for which the fundamental deviation (upper deviation) has a negative value, that is, its dimension falls below the basic size having IT9 tolerance.

Depending on the application, numerous fits ranging from extreme clearance to extreme interference can be selected using a suitable combination of fundamental deviations and fundamental tolerances. From Fig. 3.16, it can be seen that the lower deviation for the holes 'A' to 'G' is above the zero line and that for 'K' to 'ZC' is below the zero line. In addition, it can be observed that for shafts 'a' to 'g', the upper deviation falls below the zero line, and for 'k' to 'zc' it is above the zero line.

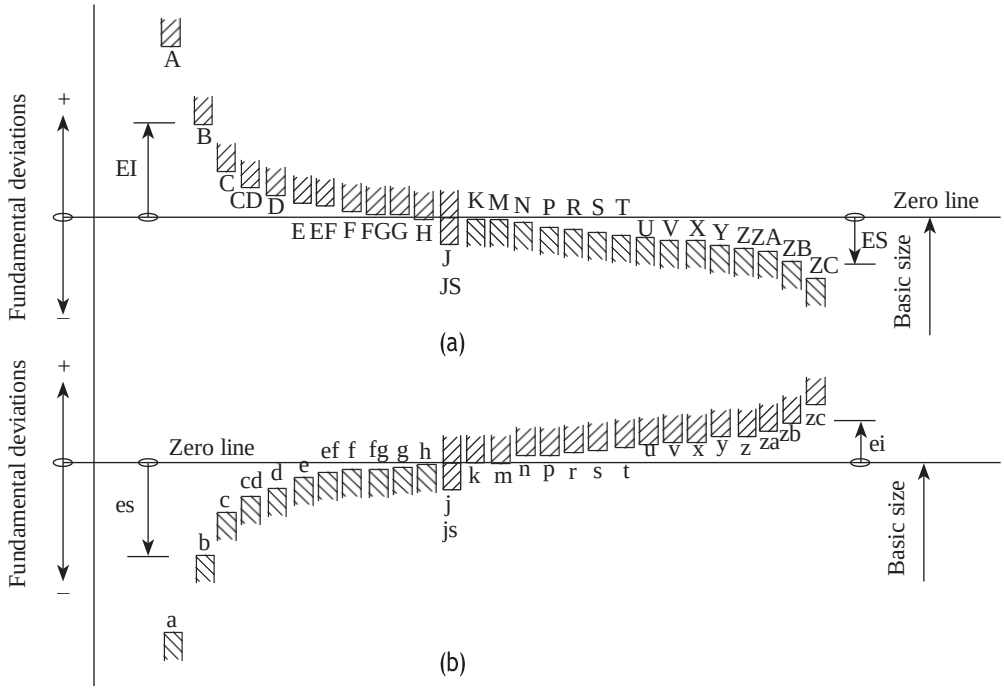


Fig. 3.16 Typical representation of different types of fundamental deviations
(a) Holes (internal features) (b) Shafts (external features)

It can be seen from Fig. 3.17 that 'EI' is above the zero line for holes 'A' to 'G', indicating positive fundamental deviation. In contrast, Fig. 3.18 shows that 'ei' is below the zero line for the shafts 'a' to 'g' and therefore the fundamental deviation is negative. In addition, from Figs 3.17 and 3.18, it can be observed that for holes 'K' to 'ZC', the fundamental deviation is negative ('EI' below the zero line), whereas for shafts 'k' to 'zc', it is positive ('ei' above the zero line).

It follows from Figs 3.17 and 3.18 that the values of 'ES' and 'EI' for the holes and 'es' and 'ei' for the shafts can be determined by adding and subtracting the fundamental tolerances, respectively. Magnitude and sign of fundamental deviations for the shafts, either upper deviation 'es' or lower deviation 'ei' for each symbol, can be determined

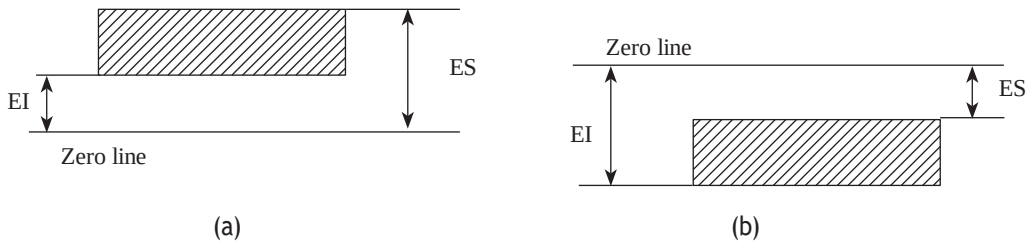


Fig. 3.17 Deviations for holes (a) Deviations for A to G (b) Deviations for K to Z_i

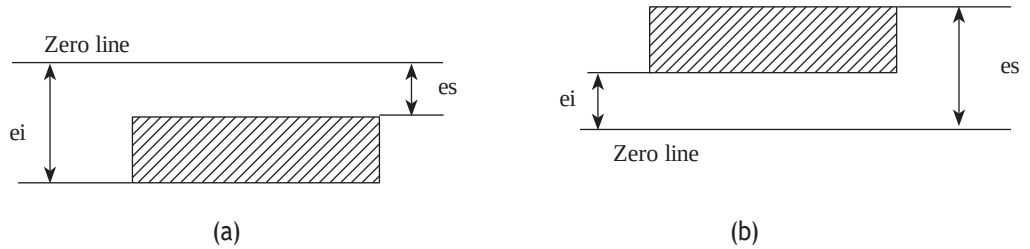


Fig. 3.18 Deviations for shafts (a) Deviations for a to g (b) Deviations for k to Z_i

from the empirical relationships listed in Tables 3.5 and 3.6, except for shafts j and js for which there is no fundamental deviation. The values given in Tables 3.5 and 3.6 are as per IS: 919. The

Table 3.5 Fundamental deviation formulae for shafts of sizes up to 500 mm

Upper deviation (es)		**Lower deviation (ei)**	
Shaft designation	In μm (for D in mm)	Shaft designation	In μm (for D in mm)
a	$= -(265 + 1.3D)$ for $D \le 120$	j5-j8	No formula
		(For js two deviations are equal to IT ± 2)	
	$= -3.5D$ for $D > 120$	k4-k7	$= +0.6\sqrt[3]{D}$
		For grades ≤k3 to ≥k8	= 0
b	$= -(140 + 0.85D)$ for $D \le 160$	m	= +(IT7 − IT6)
	$= -1.8D$ for $D > 160$	n	$= +5D^{0.34}$
		p	= + IT7 + 0 to 5
c	$= -52D^{0.2}$ for D + IT7 + 0 to 40	r	= Geometric mean of ei values for p and s
	$= -(95 + 0.8D)$ for $D > 40$		= +IT8 + 1 to 4 for $D \le 50$
d	$= -16D^{0.44}$	s	= +IT7 + 0.4D for $D > 50$
e	$= -11D^{0.41}$	t	= +IT7 + 0.63D
f	$= -5.5D^{0.41}$	u	= +IT7 + D
g	$= -2.5D^{0.34}$	v	= +IT7 + 1.25D

(Contd)

Table 3.5 (Contd)

h	= 0	x	= +IT7 + 1.6*D*
		y	= +IT7 + 2*D*
		z	= +IT7 + 2.5*D*
		za	= +IT8 + 3.15*D*
		zb	= +IT9 + 4*D*
		zc	= +IT10 + 5*D*

Table 3.6 Fundamental deviation formulae for holes of sizes up to 500 mm

For all deviations except the following:			General rule: Hole limits are identical with the shaft limits of the same symbol (letter and grade) but disposed on the other side of the zero line EI = upper deviation es of shaft of the same letter symbol but of opposite sign
For sizes above 3 mm	N	9 and coarser grades	ES = 0
	J, K, M, and N	Up to grade 8 inclusive	Special rule: ES = lower deviation ei of the shaft of the same letter symbol but one grade finer and of opposite sign, increased by the difference between the tolerances of the two grades in question
	P to ZC	Up to grade 7 inclusive	

other deviations of the holes and shafts can be determined using the following relationships that can be derived from Figs 3.17 and 3.18:

1. For holes A to G, EI is a positive fundamental deviation and EI = ES − IT.
2. For holes K to Z*i*, the fundamental deviation ES is negative and ES = EI − IT.
3. For shafts 'a' to 'g', the fundamental deviation es is negative and es = ei − IT.
4. For shafts 'k' to 'Z*i*', the fundamental deviation ei is positive, and ei = es − IT.

Production of large-sized components is associated with problems pertaining to manufacture and measurement, which do not exist in the case of smaller-sized components. As the size of the components to be manufactured increases, the difficulty in making accurate measurements also increases. Variations in temperature also affect the quality of measurement.

When the size of the components to be manufactured exceeds 500 mm, the tolerance grades IT01–IT5 are not provided, as they are considered to be too small. The fundamental tolerance unit in case of sizes exceeding 500 and up to 3150 mm is determined as follows:

$i = 0.004D + 2.1D$, where D is 0.001 mm.

The fundamental deviations for holes and shafts are given in Table 3.7, which are as per IS: 2101.

Table 3.7 Fundamental deviation for shafts and holes of sizes from above 500 to 3150 mm

Shafts			Holes			Formula for deviations in μm
Type	Fundamental deviation	Sign	Type	Fundamental deviation	Sign	(for D in mm)
d	es	-	D	EI	+	$16D^{0.44}$
e	es	-	E	EI	+	$11D^{0.41}$
f	es	-	F	EI	+	$5.5D^{0.41}$
g	es	-	G	EI	+	$2.5D^{0.34}$
h	es	No sign	H	EI	No sign	0
js	ei	-	JS	ES	+	$0.5IT\pi$
k	ei	-	K	ES	-	0
m	ei	+	M	ES	-	$0.024D + 12.6$
n	ei	+	N	ES	-	$0.04D + 21$
P	ei	+	P	ES	-	$0.072D + 37.8$
r	ei	+	R	ES	-	Geometric mean of the values for p and s or P and S
s	ei	+	S	ES	-	$IT7 + 0.4D$
t	ei	+	T	ES	-	$IT7 + 0.63D$
u	ei	+	U	ES	-	$IT7 + D$

3.6.1 General Terminology

The following are the commonly used terms in the system of limits and fits.

Basic size This is the size in relation to which all limits of size are derived. Basic or nominal size is defined as the size based on which the dimensional deviations are given. This is, in general, the same for both components.

Limits of size These are the maximum and minimum permissible sizes acceptable for a specific dimension. The operator is expected to manufacture the component within these limits. The maximum limit of size is the greater of the two limits of size, whereas the minimum limit of size is the smaller of the two.

Tolerance This is the total permissible variation in the size of a dimension, that is, the difference between the maximum and minimum limits of size. It is always positive.

Allowance It is the intentional difference between the LLH and HLS. An allowance may be either positive or negative.

$$\text{Allowance} = \text{LLH} - \text{HLS}$$

Grade This is an indication of the tolerance magnitude; the lower the grade, the finer the tolerance.

Deviation It is the algebraic difference between a size and its corresponding basic size. It may be positive, negative, or zero.

Upper deviation It is the algebraic difference between the maximum limit of size and its corresponding basic size. This is designated as 'ES' for a hole and as 'es' for a shaft.

Lower deviation It is the algebraic difference between the minimum limit of size and its corresponding basic size. This is designated as 'EI' for a hole and as 'ei' for a shaft.

Actual deviation It is the algebraic difference between the actual size and its corresponding basic size.

Fundamental deviation It is the *minimum* difference between the size of a component and its basic size. This is identical to the upper deviation for shafts and lower deviation for holes. It is the closest deviation to the basic size. The fundamental deviation for holes are designated by capital letters, that is, A, B, C, …, H, …, ZC, whereas those for shafts are designated by small letters, that is, a, b, c…, h…, zc. The relationship between fundamental, upper, and lower deviations is schematically represented in Fig. 3.19.

Zero line This line is also known as the line of zero deviation. The convention is to draw the zero line horizontally with positive deviations represented above and negative deviations indicated below. The zero line represents the basic size in the graphical representation.

Shaft and hole These terms are used to designate all the external and internal features of any shape and not necessarily cylindrical.

Fit It is the relationship that exists between two mating parts, a hole and a shaft, with respect to their dimensional difference before assembly.

Maximum metal condition This is the maximum limit of an external feature; for example, a shaft manufactured to its high limits will contain the maximum amount of metal. It is also the minimum limit of an internal feature; for example, a component that has a hole bored in it to its lower limit of size will have the minimum amount of metal removed and remain in its

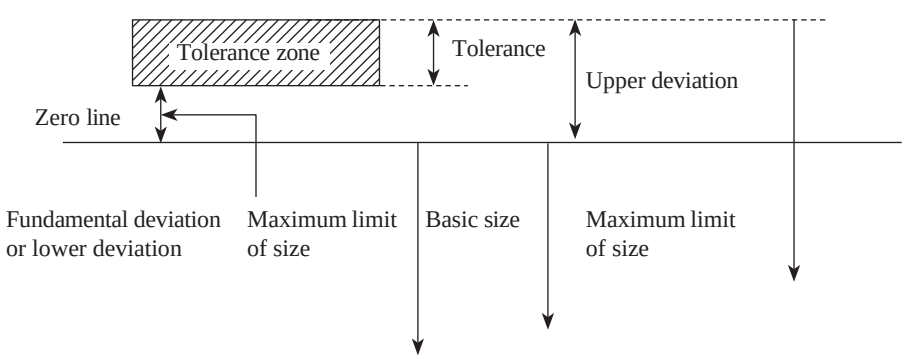


Fig. 3.19 Relationship between fundamental, upper, and lower deviations

maximum metal condition, (i.e., this condition corresponds to either the largest shaft or the smallest hole). This is also referred to as the GO limit.

Least metal condition This is the minimum limit of an external feature; for example, a shaft will contain minimum amount of material, when manufactured to its low limits. It is also the maximum limit of an internal feature; for example, a component will have the maximum amount of metal removed when a hole is bored in it to its higher limit of size, this condition corresponds to either the smallest shaft or the largest hole. This is also referred to as the NO GO limit.

Tolerance zone The tolerance that is bound by the two limits of size of the component is called the tolerance zone. It refers to the relationship of tolerance to basic size.

International tolerance grade (IT) Tolerance grades are an indication of the degree of accuracy of the manufacture. Standard tolerance grades are designated by the letter IT followed by a number, for example, IT7. These are a set of tolerances that varies according to the basic size and provides a uniform level of accuracy within the grade.

Tolerance class It is designated by the letter(s) representing the fundamental deviation followed by the number representing the standard tolerance grade. When the tolerance grade is associated with letter(s) representing a fundamental deviation to form a tolerance class, the letters IT are omitted and the class is represented as H8, f7, etc.

Tolerance symbols These are used to specify the tolerance and fits for mating components. For example, in 40 H8f7, the number 40 indicates the basic size in millimetres; capital letter H indicates the fundamental deviation for the hole; and lower-case letter f indicates the shaft. The numbers following the letters indicate corresponding IT grades.

3.6.2 Limit Gauging

Eli Whitney, who is hailed as the father of the American system, won the first contract in 1798 for the production of muskets and developed the gauging system. Richard Roberts of the United Kingdom first used plug and collar gauges for dimensional control. In 1857, Joseph Whitworth demonstrated internal and external gauges for use with a shaft-based limit system.

As discussed in Section 3.1, in mass production, components are manufactured in accordance with the permissive tolerance limits, as suggested by the designer. Production of components within the permissive tolerance limits facilitates interchangeable manufacture. It is also essential to check whether the dimensions of the manufactured components are in accordance with the specifications or not. Therefore, it is required to control the dimensions of the components. Several methods are available to achieve the control on dimensions. Various precision measuring instruments can be used to measure the actual dimensions of the components, which can be compared with the standard specified dimensions to decide the acceptability of these components.

In mass production, where a large number of similar components are manufactured on an interchangeable basis, measuring the dimensions of each part will be a time-consuming and expensive exercise. In addition, the actual or absolute size of a component, provided that it is within the limits specified, is not of much importance because the permissible limits of

variations in dimensions would have been specified by the designer in a logical manner, giving due consideration to the functional requirements. Therefore, in mass production, gauges can be used to check for the compliance of the limits of the part with the permissive tolerance limits, instead of measuring the actual dimensions. The term 'limit gauging' signifies the use of gauges for checking the limits of the components. Gauging plays an important role in the control of dimensions and interchangeable manufacture.

Limit gauges ensure that the components lie within the permissible limits, but they do not determine the actual size or dimensions. Gauges are scaleless inspection tools, which are used to check the conformance of the parts along with their forms and relative positions of the surfaces of the parts to the limits. The gauges required to check the dimensions of the components correspond to two sizes conforming to the maximum and minimum limits of the components. They are called GO gauges or NO GO or NOT GO gauges, which correspond, respectively, to the MML and LML of the component, as depicted in Figs 3.20 and 3.21. As discussed in section 3.4, MML is the lower limit of a hole and higher limit of the shaft and LML corresponds to the higher limit of a hole and lower limit of the shaft. The GO gauge manufactured to the maximum limit will assemble with the mating (opposed) part, whereas the NOT GO gauge corresponding to the low limit will not, hence the names GO and NOT GO gauges, respectively.

Practically, every gauge is a replica of the part that mates with the part for which the gauge has been designed. Consider an example of the manufacture of a cylinder that mates with a piston. The plug gauge, using which the cylinder bore is checked, is a copy of the opposed part (piston) as far as its form and size are concerned. When a gauge is designed as a replica

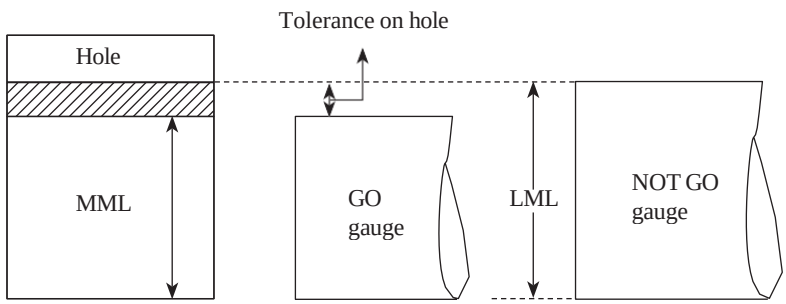


Fig. 3.20 Metal limits for hole gauging

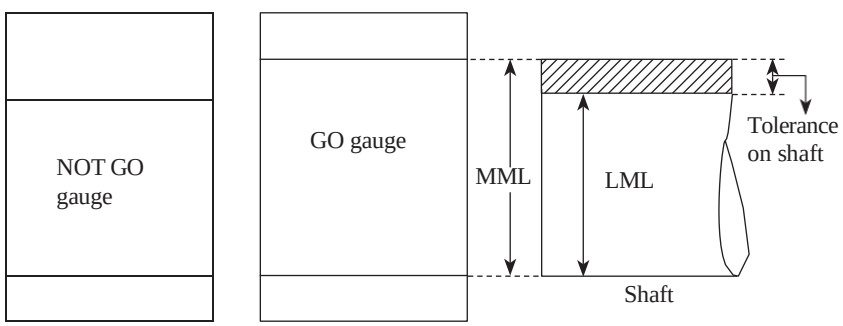


Fig. 3.21 Metal limits for shaft gauging

of the mating (opposed) part so far as the dimension to be checked is concerned, it is known as 'standard gauge'. The main intention in the design of gauges is simplicity, which helps in making continuous and accurate measurements.

It is important to note that normally clearance fits are preferred for a majority of the assembly operations. Allowance or minimum clearance is determined by the algebraic difference of the MMLs of the mating parts. Therefore, for clearance fits, the MMLs of the mating parts become more critical than the LMLs. This assumes importance for the following two reasons:

1. MMLs are crucial for effective functioning of the parts.
2. If the MMLs slightly exceed the specified values then assembly itself becomes impossible.

As discussed earlier, for gauging the MMLs of the mating parts, GO gauges are used. Therefore, it becomes imperative that special attention needs to be given when GO gauges are designed for gauging these limits. Whenever the components are gauged for their MMLs, if the GO gauges fail to assemble during inspection, the components should not be accepted under any circumstances. The minimum limits in a clearance fit of a product are not so critical because even if they exceed the specified limits and the NOT GO gauge assembles, its acceptance may result in functional degradation and because of the reduced quality the useful life of the product may get affected. Hence, it becomes essential that more care is taken especially when GO gauges are used, when compared to NOT GO gauges during inspection.

3.6.3 Classification of Gauges

The detailed classification of the gauges is as follows:

1. Plain gauges

(a) According to their type:
   (i) Standard gauges
   (ii) Limit gauges

(b) According to their purpose:
   (i) Workshop
   (ii) Inspection
   (iii) Reference, or master, or control gauges

(c) According to the form of the tested surface:
   (i) Plug gauges for checking holes
   (ii) Snap and ring gauges for checking shafts

(d) According to their design:
   (i) Single- and double-limit gauges
   (ii) Single- and double-ended gauges
   (iii) Fixed and adjustable gauges

2. Adjustable-type gap gauges
3. Miscellaneous gauges
   (a) Combined-limit gauges
   (b) Taper gauges
   (c) Position gauges
   (d) Receiver gauges

(e) Contour gauges
(f) Profile gauges

3.6.4 Taylor's Principle

In 1905, William Taylor developed a concept relating to the gauging of components, which has been widely used since then. Since World War II, the term Taylor's principle has generally been applied to the principle of limit gauging and extensively used in the design of limit gauges. Prior to 1905, simple GO gauges were used. The components were carefully manufactured to fit the gauges. Since NOT GO gauges were not used, these components were without tolerance on their dimensions.

The theory proposed by Taylor, which is extensively used in the design of limit gauges, not only defines the function, but also defines the form of most limit gauges.

Taylor's principle states that the GO gauge is designed to check maximum metal conditions, that is, LLH and HLS. It should also simultaneously check as many related dimensions, such as roundness, size, and location, as possible.

The NOT GO gauge is designed to check minimum metal conditions, that is, HLH and LLS. It should check only one dimension at a time. Thus, a separate NOT GO gauge is required for each individual dimension.

During inspection, the GO side of the gauge should enter the hole or just pass over the shaft under the weight of the gauge without using undue force. The NOT GO side should not enter or pass.

The basic or nominal size of the GO side of the gauge conforms to the LLH or HLS, since it is designed to check maximum metal conditions. In contrast, the basic or nominal size of the NOT GO gauge corresponds to HLH or LLS, as it is designed to check minimum metal conditions.

It can be seen from Fig. 3.22 that the size of the GO plug gauge corresponds to the LLH and the NOT GO plug gauge to the HLH. Conversely, it can be observed from Fig. 3.23 that the GO snap gauge represents the HLS, whereas the NOT GO snap gauge represents the LLS.

It is pertinent to discuss here that since the GO plug is used to check more than one dimension of the hole simultaneously, the GO plug gauge must have a full circular section and must be of full length of the hole so that straightness of the hole can also be checked. During inspection, it can be ensured that if there is any lack of straightness or roundness of the hole a full entry of the GO plug gauge will not be allowed. Thus, it not only controls the diameter in any given cross-section but also ensures better bore alignment. However, it should be mentioned here that the

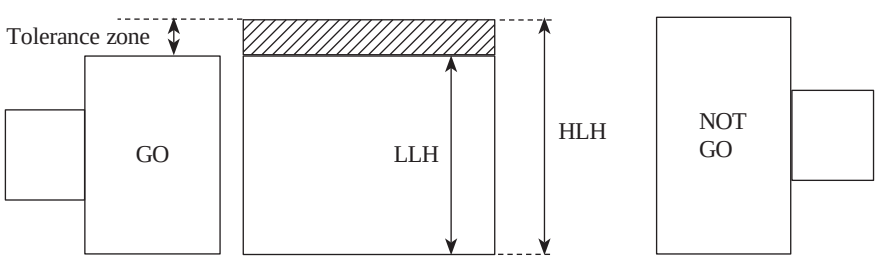


Fig. 3.22 GO and NOT GO limits of plug gauge

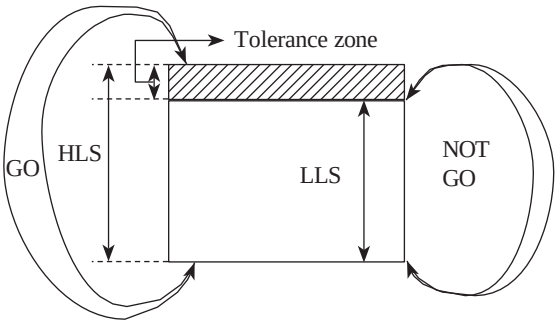


Fig. 3.23 GO and NOT GO limits of snap gauge

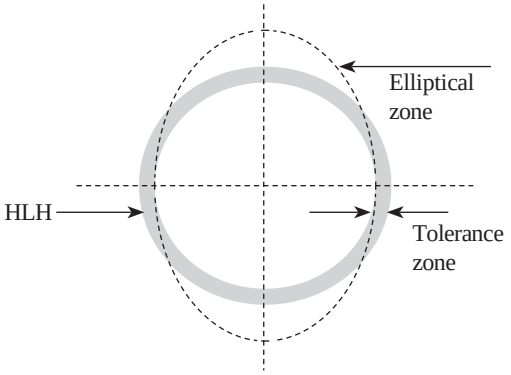


Fig. 3.24 Ovality in hole

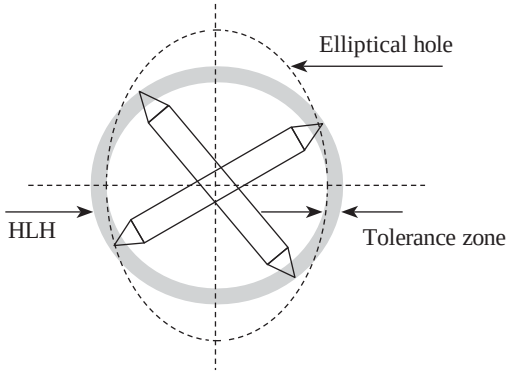


Fig. 3.25 Pin-type NOT GO to the check ovality

GO plug gauge cannot check the degree of ovality. The short GO plug gauge, if used in inspection, will pass through all the curves and is hence not possible to identify defective parts. Therefore, in order to get good results, this condition has to be fulfilled during the inspection of the parts. The length of the plug should normally be more than 1.5 times the diameter of the hole to be checked. Compared to GO plug gauges, the NOT GO plug gauges are relatively shorter.

Let us consider the gauging of a cylindrical hole. A simple plug gauge is used to gauge this hole. During inspection, the GO gauge, which measures the minimum limit of the hole, enters but the NOT GO gauge, which corresponds to the maximum limit, does not. In this case, according to Taylor's theorem, the hole is considered to be within the specified limits and hence accepted. However, the shape of the hole has not been taken into account here. Most of the methods used in the manufacture of the holes are capable of producing truly circular holes. When these holes are gauged, they are accepted if they are within the specified tolerance zone. As long as there are no circularity errors, there is no problem. However, when the holes deviate from circularity, the problem starts. Consider an elliptical hole as shown in Fig. 3.24. It can be clearly observed from Fig. 3.24 that the NOT GO gauge does not enter because the minor axis of the ellipse is smaller than the HLH, which corresponds to the diameter of the NOT GO gauge. Therefore, even if the hole is slightly elliptical, the gauge does not take into account this variation of the hole shape and it still gets accepted as the GO gauge assembles and the NOT GO gauge does not.

One way of overcoming this problem is to employ the NOT GO, which is in the form of a pin. Any error in the circularity can easily be detected by placing the pin at different cross-sections of the hole, as depicted in Fig. 3.25. Hence, it can be said that Taylor's principle does not take care of the error of form, circularity, or straightness, and some modifications are needed to overcome these limitations.

3.6.5 Important Points for Gauge Design

The following points must be kept in mind while designing gauges:

1. The form of GO gauges should be a replica of the form of the opposed (mating) parts.

2. GO gauges enable several related dimensions to be checked simultaneously and hence are termed complex gauges.
3. During inspection, GO gauges must always be put into conditions of maximum impassability.
4. NOT GO gauges check a single element of feature at a time.
5. In inspection, NOT GO gauges must always be put into conditions of maximum passability.

3.6.6 Material for Gauges

The material used to manufacture gauges should satisfy the following requirements:

1. The material used in the manufacture of gauges should be hard and wear resistant for a prolonged life. This is one of the most important criteria that should be fulfilled.
2. It should be capable of maintaining dimensional stability and form.
3. It should be corrosion resistant.
4. It should be easily machinable, in order to obtain the required degree of accuracy and surface finish.
5. It should have a low coefficient of expansion, in order to avoid temperature effect.

High-carbon steel is the most suitable and inexpensive material used for manufacturing gauges. It can be heat treated suitably to provide stability and high hardness. It can easily be machined to a high degree of accuracy.

Mild steel gauges are the most suitable for larger gauges. They are suitably heat treated by carburizing to the required depth and then case hardened on their working surfaces to allow for final grinding and finishing. After hardening, internal stresses are relieved to improve stability.

Chromium-plated gauges are very popular and extensively used for gauging. Chromium plating makes the surface of the gauge very hard, and resistant to abrasion and corrosion. It is also very useful in reclaiming worn-out gauges. For gauging aluminium or other materials having an abrasive action, chromium-plated gauges are extensively used. The plug gauges employed for gauging have satellite ribs that are inserted in the gauging surface.

Glass gauges are not very popular although they have good wear and corrosion resistance properties. The problem with these gauges is that they either get damaged or are easily broken if dropped. They are not affected by changes in temperature and have very low coefficient of thermal expansion.

Although Invar, which contains 36% of nickel, has a low coefficient of expansion, it is not suitable over a long period. Elinvar has 42% of nickel, is more stable than Invar, and also has a low coefficient of expansion.

3.6.7 Gauge Tolerance (Gauge Maker's Tolerance)

We know that gauges have to be manufactured to their required dimensions corresponding to their maximum metal conditions. Gauges, like any other component, cannot be manufactured to their exact size or dimensions. In order to accommodate these dimensional variations, which arise due to the limitations of the manufacturing process, skill of the operator, etc., some tolerance must be allowed in the manufacture of gauges. Thus, the tolerance that is allowed in the manufacture of gauges is termed *gauge maker's tolerance* or simply *gauge tolerance.* Logically, gauge tolerance should be kept as minimum as possible; however, this increases the gauge manufacturing cost. There is no universally accepted policy for deciding the amount of

tolerance to be provided on gauges. The normal practice is to take gauge tolerance as 10% of the work tolerance.

3.6.8 Wear Allowance

As discussed in Section 3.6.4, according to Taylor's principle, during inspection the NOT GO side should not enter or pass. The NOT GO gauge seldom engages fully with the work and therefore does not undergo any wear. Hence, there is no need to provide an allowance for wear in case of NOT GO gauges.

Taylor's principle also states that the GO side of the gauge should enter the hole or just pass over the shaft under the weight of the gauge without using undue force. During inspection, the measuring surfaces of the gauge constantly rub against the mating surfaces of the workpiece. Therefore, the GO gauges suffer wear on the measuring surfaces and thus lose their initial dimension. Hence, wear allowance is provided for GO gauges to extend their service life. As a consequence of this wear, the size of the GO plug gauge decreases while that of the ring or gap gauge increases. The wear allowance provided for the GO gauges are added in a direction opposite to wear. This allowance is added in for a plug gauge while subtracted for a ring or gap gauge. A wear allowance of 10% of gauge tolerance is widely accepted in industries. If the work tolerance of a component is less than 0.1 mm, no wear allowance on gauges is provided for that component, since a wear allowance of less than 0.001 mm will not have any practical effect on the gauges.

The allowance on new gauge is made by fixing the tolerance zone for the gauge from the MML of the work by an amount equal to the wear allowance. A new gauge is then manufactured within the limits specified by the tolerance zone for the gauge in this position. The gauge is then allowed to wear with use until its size coincides with the maximum material limit for the work.

3.6.9 Methods of Tolerance Specification on Gauges

By considering all the important points of gauge design, tolerances on gauges (snap and plug gauges) are specified by the following three methods.

First Method

This method of specifying tolerances on gauges was evolved during the developmental stages of gauges. In this system, the tolerance zones of workshop and inspection gauges are different and are made separately. Here the tolerances on the workshop gauge are set to fall inside the work tolerance, whereas the inspection gauge tolerances fall outside the work tolerance.

In workshop gauges, the tolerance on GO and NOT GO gauges is 10% of the work tolerance. For example, if the work tolerance is 10 units, then only 8 units will be left as the difference between maximum limit of GO gauge and minimum limit of NOT GO gauge, since these tolerances fall inside and are 1 unit each, as shown in Fig. 3.26. For inspection gauges, it can be seen that the tolerance on gauges is kept beyond 10% of work tolerance.

Disadvantages of workshop and inspection gauges:

1. Although some of the components are well within the work tolerance limits, they may be rejected under workshop gauges. These components again need to be checked by inspection gauges and may be accepted after that.

2. It may also happen that some of the components when tested by inspection gauges may get accepted.
3. As the tolerance zones are different, both workshop and inspection gauges have to be manufactured separately.

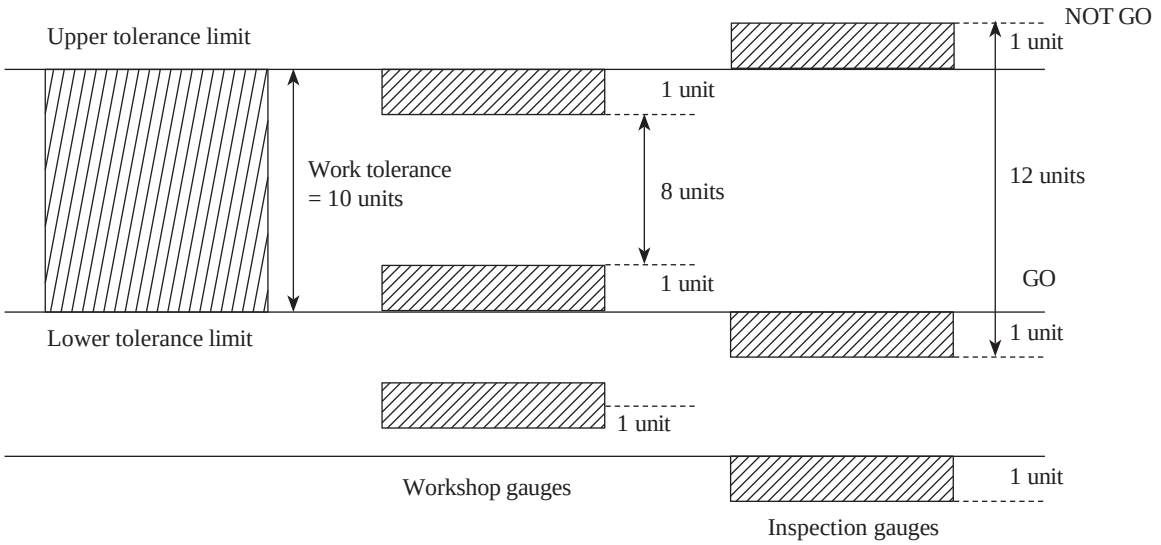


Fig. 3.26 Disposition of tolerances on workshop and inspection gauges

Revised System of Gauges

In this system, the tolerance zone on inspection gauges is reduced and the tolerance on workshop gauges is retained as before. The disadvantages of inspection gauges are reduced due to the reduction of tolerance on these gauges. It can be observed from Fig. 3.27 that instead of 120% of the tolerance provided in the first system for GO and NOT GO inspection gauges, in the revised system, 110% of the range of work tolerance is provided.

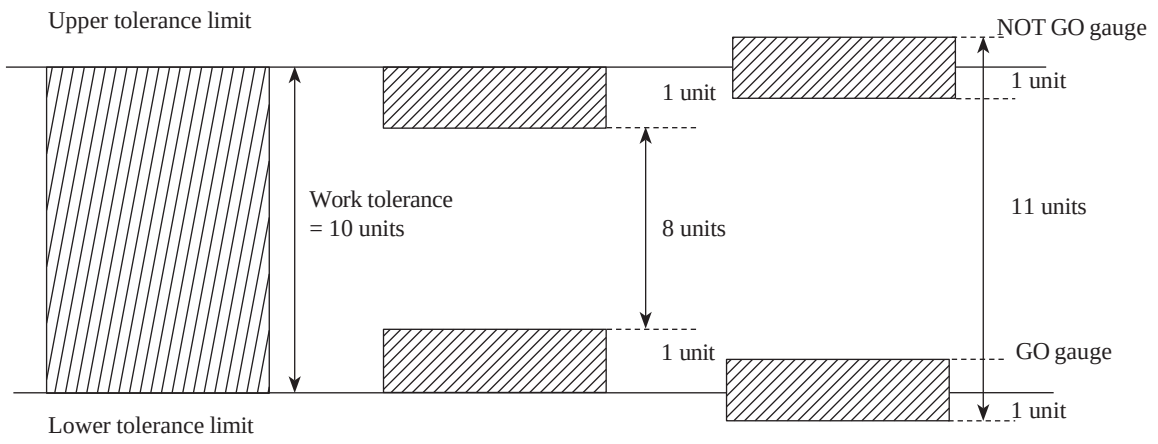


Fig. 3.27 Disposition of modified tolerances on workshop and inspection gauges

British System

The present British system is in accordance with Taylor's principle and is widely accepted in industries. This system works according to the following guidelines:

1. The specified tolerance should be wide and, at the same time, consistent with satisfactory functioning, economical production, and inspection.
2. During inspection, the work that lies outside the specified limits should not be accepted.

Thus, in this modern system, the same tolerance limits are specified on workshop and inspection gauges, and the same gauges can be used for both purposes.

It can be observed from Fig. 3.28 that the tolerance zone for the GO gauges is placed inside the work limits and that for the NOT GO gauges is outside the work limits.

A margin is provided in between the tolerance zone for the gauge and MML of the work to accommodate wear allowance for GO gauges. It is to be noted here that wear allowance should not be permitted beyond the MML of the work, especially when the limit is of critical importance. The magnitude of wear allowance is 10% of the gauge tolerance.

3.6.10 Numerical Examples

Example 3.7 Design the general type of GO and NOT GO gauges as per the present British system for a 40 mm shaft and hole pair designated as 40 H8/d9, given that

(a) $i = 0.453\sqrt[3]{D} + 0.001D$
(b) 40 mm lies in the diameter range of 30–50 mm
(c) IT8 = 25i
(d) IT9 = 40i
(e) upper deviation of shaft = $-16D^{0.44}$
(f) wear allowance assumed to be 10% of gauge tolerance

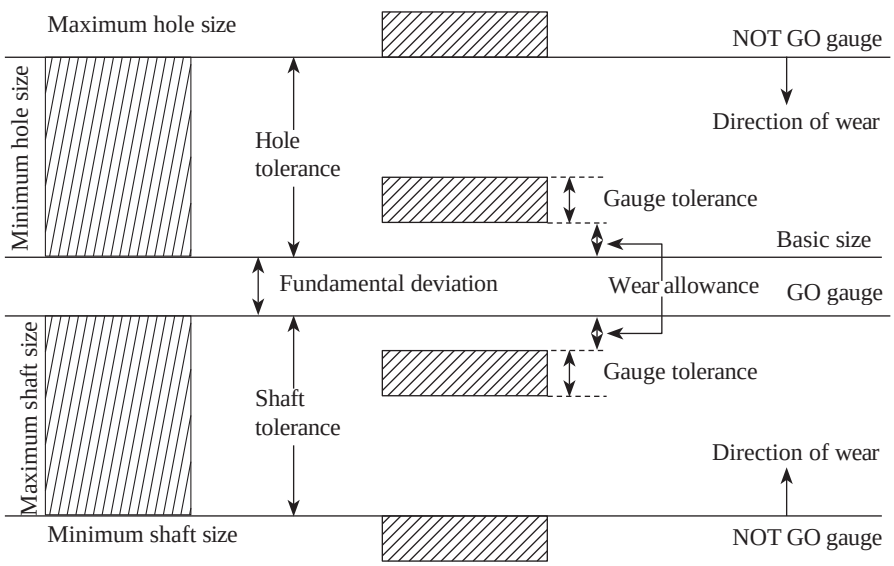


Fig. 3.28 Disposition of tolerances and allowances on gauges

Solution

The standard diameter steps for 40 mm shaft falls in the range of 30–50 mm.

D can be calculated using the equation $\sqrt{D_{max} \times D_{min}}$.
Therefore,

$$D = \sqrt{30 \times 50}$$
$$D = 38.7298 \text{ mm}$$

The value of fundamental tolerance unit is given by

$$i = 0.453 \sqrt[3]{D} + 0.001D$$

$$i = 0.453 \left(\sqrt[3]{38.7298} \right) + 0.001(38.7298) = 1.571 \text{ µm}$$

For hole quality H8, the fundamental tolerance is 25i.

$$25i = 25(1.571) = 39.275 \text{ µm} = 0.039275 \text{ mm} \approx 0.039 \text{ mm}$$

For hole, the fundamental deviation is zero.
Hence, hole limits are as follows:

LLH = 40 mm
HLH = 40.00 + 0.039 = 40.039 mm
Hole tolerance = 40.039 − 40 = 0.039 mm

For shaft quality d9, the fundamental tolerance is 40i.

$$40i = 40(1.571) = 62.84 \text{ µm} = 0.06284 \text{ mm} \approx 0.063 \text{ mm}$$

For d shaft, the fundamental deviation is given by $-16D^{0.44}$

Therefore, fundamental deviation $= -16(38.7298)^{0.44}$
$= -79.9576 \text{ µm} \approx -0.07996 \text{ mm} \approx -0.080 \text{ mm}$

Hence, shaft limits are as follows:

HLS = 40 − 0.080 = 39.92 mm
LLS = 40 − (0.080 + 0.063) = 39.857 mm
Shaft tolerance = 39.92 − 39.857 = 0.063 mm

Hence, the hole and shaft limits are as follows:

$$\text{Hole} = 40 \, ^{+0.039}_{+0.000} \text{ mm and shaft} = 40 \, ^{-0.080}_{-0.143} \text{ mm}$$

The disposition of tolerances is as shown in Fig. 3.29.

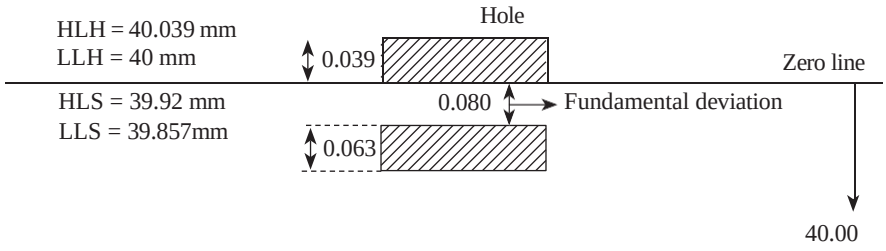


Fig. 3.29 Disposition of tolerances

CHAPTER-2 Linear Measurement

After studying this chapter, the reader will be able to

- understand the basic principles of design of linear measuring instruments
- explain the use of 'datum planes' in dimensional measurement
- appreciate the advantages offered by scaled instruments in contrast to a simple steel rule, and discuss the various applications and limitations of their variants
- elucidate the vernier family of instruments for linear measurement and read vernier instruments
- describe how Abbe's law applies to micrometer measurement
- explain how the micrometer principle has developed into a family of diverse instruments
- utilize the digital electronic instruments in linear measurement
- explain the use of slip gauges, and their manufacture and calibration

4.1 INTRODUCTION

In Chapter 2, the methods by which engineering standards of length are established were discussed. Both direct and indirect linear measuring instruments conform to these established standards of length and provide convenient means for making accurate and precise linear measurements. Vernier calliper and vernier micrometer are the most widely used linear measuring instruments in machine shops and tool rooms. Measuring instruments are designed either for line measurements (e.g., steel rule or vernier calliper) or for end measurements in order to measure the distance between two surfaces using an instrument (e.g., screw gauge). Callipers and dividers, which are also linear measurement devices, are basically *dimension transfer instruments*. They will not directly provide the measurement of length on a scale. Quality of measurement not only depends on the accuracy of these instruments, but also calls for application of certain simple principles to be followed during measurements. Illustrations are given throughout this chapter, especially on the latter issue, to highlight that care should be exercised for the proper use of linear measuring instruments.

Most people's first contact with linear measurement is with a steel rule or a tape measure. However, today's engineer has a choice of a wide range of instruments—from purely

mechanically operated instruments to digital electronics instruments. One has to consider only the nature of application and cost of measurement to decide which instrument is the best for an application. This chapter covers a broad range of linear measurement instruments, from a simple steel rule to digital callipers and micrometers. However, many of these instruments, such as depth gauge and height gauge, need to be used with a datum to ensure accuracy of measurements. The foundation for all dimensional measurements is the 'datum plane', the most important ones being the surface plate and the V-block. Constructions of the surface plate and V-block are also explained with illustrations.

4.2 DESIGN OF LINEAR MEASUREMENT INSTRUMENTS

The modern industry demands manufacture of components and products to a high degree of dimensional accuracy and surface quality. Linear measurement instruments have to be designed to meet stringent demands of accuracy and precision. At the same time, the instruments should be simple to operate and low priced to make economic sense for the user. Proper attachments need to be provided to make the instrument versatile to capture dimensions from a wide range of components, irrespective of the variations in cross-sections and shapes. The following points highlight important considerations that have to be addressed in the design of linear measurement instruments:

1. The measuring accuracy of line-graduated instruments depends on the original accuracy of the line graduations. Excessive thickness or poor definition of graduated lines affects the accuracy of readings captured from the instrument.
2. Any instrument incorporating a scale is a suspect unless it provides compensation against wear.
3. Attachments can enhance the versatility of instruments. However, every attachment used along with an instrument, unless properly deployed, may contribute to accumulated error. Wear and tear of attachments can also contribute to errors. Use attachments when their presence improves reliability more than their added chance for errors decreasing it.
4. Instruments such as callipers depend on the feel of the user for their precision. Good quality of the instrument promotes reliability, but it is ultimately the skill of the user that ensures accuracy. Therefore, it is needless to say that proper training should be imparted to the user to ensure accurate measurements.
5. The principle of alignment states that the line of measurement and the line of dimension being measured should be coincident. This principle is fundamental to good design and ensures accuracy and reliability of measurements.
6. Dial versions of instruments add convenience to reading. Electronic versions provide digital readouts that are even easier to read. However, neither of these guarantees accuracy and reliability of measurements unless basic principles are adhered to.
7. One important element of reliability of an instrument is its *readability*. For instance, the smallest division on a micrometer is several times larger than that on a steel rule of say 0.1 mm resolution, which is difficult to read. However, the micrometer provides better least count, say up to 0.01 mm, compared to the same steel rule. Therefore, all other things being equal, a micrometer is more reliable than even a vernier scale. However, micrometers have a lesser range than verniers.

8. If cost is not an issue, digital instruments may be preferred. The chief advantage of the electronic method is the ease of signal processing. Readings may be directly expressed in the required form without additional arithmetic. For example, they may be expressed in either metric or British units, and can also be stored on a memory device for further use and analysis.
9. Whenever a contact between the instrument and the surface of the job being measured is inevitable, the contact force should be optimum to avoid distortion. The designer cannot leave the fate of the instrument on the skill of the user alone. A proper device like a *ratchet stop* can limit the contact force applied on the job during measurements, thereby avoiding stress on the instrument as well as distortion of the job.

4.3 SURFACE PLATE

In Section 4.2, we understood that every linear measurement starts at a reference point and ends at a measured point. This is true when our basic interest is in measuring a single dimension, length in this case. However, the foundation for all dimensional measurements is the 'datum plane', the most important one being the surface plate. A surface plate is a hard, solid, and horizontal flat plate, which is used as the reference plane for precision inspection, marking out, and precision tooling set-up (Fig. 4.1). Since a surface plate is used as the datum for all measurements on a job, it should be finished to a high degree of accuracy. It should also be robust to withstand repeated contacts with metallic workpieces and not be vulnerable to wear and tear.

The history of surface plates can be traced to the early 19th century when Richard Robert invented the planer in 1817, which was presumably the first machine tool that used a flat surface. He showed a way of duplicating flat surfaces with a high degree of accuracy, and the world of sliding motions and flat surfaces was born. However, the surface plates used by Roberts were of quite low accuracy compared to today's standards. One should credit the contribution of Sir Joseph Whitworth, a leading name in metrology, who recognized the lack of understanding of the concept of flatness at that time and devised a methodology in the year 1840 for generating a flat surface by the 'three-plate method'. This method is being used even today to manufacture surface plates, although better and modern methods of fabricating surface plates are becoming increasingly popular. In this method, three cast iron plates with ribbed construction (for rigidity) are rough machined along their edges and top surfaces. The plates are kept in the open for normalizing for about a year. Natural changes in temperature relieve the internal stresses. The plates are then finish-machined to a high degree of accuracy and are marked #1, #2, and #3, and applied with a coating of Prussian blue. In a six-step process, the surfaces of two of the plates are placed in contact in a particular order and the blued portions are scraped. The pairing of the plates is varied in a pre-planned sequence, which ensures that all three surfaces match to a high degree, thereby ensuring accurate flat surfaces.

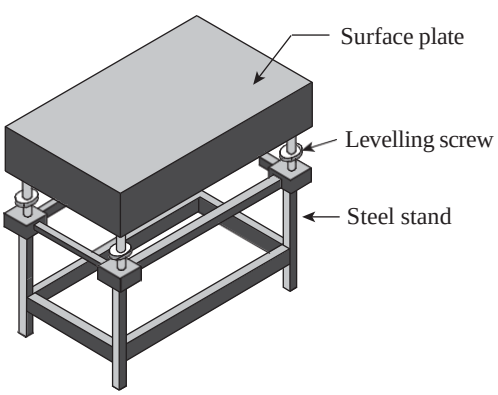


Fig. 4.1 Surface plate

The surface plates are made either from cast iron or from granite. Even though granite surface plates are perceived to be superior, cast iron surface plates are still in wide use. In fact, a cast iron surface plates is used as a tool for lapping granite surface plates to the required degree of accuracy. Cast iron allows itself to be impregnated with the lapping media over a large flat surface. In the following paragraphs, we will look into the construction and use of cast iron and granite surface plates in greater detail.

Cast Iron Surface Plates

Despite a drop in their usage, cast iron surface plates still retain popularity as surface masters. They are made of either plain or alloyed close-grained cast iron, reinforced with ribs to provide strength against bending or buckling. IS2285-1991 specifies the composition, size, and cross-sectional details of ribs and thickness of plates. The plates are manufactured in three grades, namely grade 0, grade I, and grade II. While grade 0 and grade I plates are hand scraped to achieve the required degree of flatness, grade II plates are precision machined to the required degree of accuracy. Table 4.1 illustrates some of the standard sizes of surface plates as per IS 2285-1991.

Table 4.1 Cast iron surface plate specifications as per IS 2285-1991

Size (mm)	Maximum deviation from flatness in microns			Approximate weight (kg)
	Grade 0	Grade I	Grade II	
300 × 300	4	7	15	21
400 × 400	4.5	9	17	50
450 × 300	4	8	16	39
450 × 450	4.5	9	18	62
600 × 450	5	10	20	79
630 × 400	5	10	20	96
600 × 600	5	10	20	128
630 × 630	5	10	21	156
900 × 600	6	12	23	204
1500 × 1200	8	16	33	986

Smaller-size plates are usually provided with a handle. All surface plates need to be provided with protective covers when not in use. Fabricated heavy angle iron stands with levelling screws provide convenient working height for surface plates. As per the proven practice devised by Sir Whitworth, surface plates are fabricated in sets of three. Cast iron is dimensionally more stable over time compared to granite plates. Unlike granite, it also has uniform optical properties with very small light penetration depth, which makes it a favourable material for certain optical applications. One significant drawback of cast iron is its higher coefficient of thermal expansion, which makes it unsuitable for applications involving large variations in temperature.

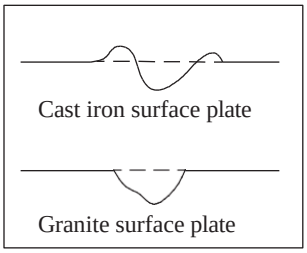


Fig. 4.2 Comparison between CI and granite surfaces

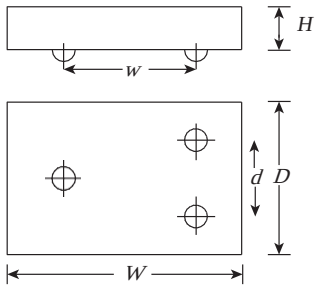


Fig. 4.3 Dimension parameters used in Table 4.2

Granite Surface Plates

In recent times, granite has replaced cast iron as the preferred material for surface plates. Most surface plates are made of black granite, while pink granite is the next preferred choice. Granite has many advantages over cast iron. Natural granite that is seasoned in the open for thousands of years is free from warp age or deterioration. It is twice as hard as cast iron and not affected by temperature changes. It is not vulnerable to rusting and is non-magnetic. As shown in Fig. 4.2, it is free from burrs or protrusions because of its very fine grain structure.

Table 4.2 lists some of the standard sizes of surface plates as per IS 2285-1991. Figure 4.3 illustrates the meaning of dimension parameters used in Table 4.2. The plates are manufactured in three grades, namely grade 00, grade 0, and grade 1. Flatness for grade 00 surface plates range from 0.002 to 0.007 mm, whereas that for grade 0 plates varies from 0.003 to 0.014 mm. Grade 1 plate is the coarsest, with flatness ranging from 0.005 to 0.027 mm.

Table 4.2 Sizes of granite surface plates as per IS 2285-1991

Grade	S. no.	Dimensions $W \times D \times H$ (mm)	Flatness (mm)	w (mm)	d (mm)	Mass (kg)
00	1	300 × 300 × 100	0.002	240	240	27
	2	450 × 300 × 100	0.002	390	240	40
	3	750 × 500 × 130	0.003	630	420	146
	4	1500 × 1000 × 200	0.004	1100	700	900
	5	3000 × 2000 × 500	0.007	2000	1500	9000
0	1	300 × 300 × 100	0.003	240	240	27
	2	450 × 300 × 100	0.003	390	240	40
	3	750 × 500 × 130	0.005	630	420	146
	4	1500 × 1000 × 200	0.008	1100	700	900
	5	3000 × 2000 × 500	0.014	2000	1500	9000
1	1	300 × 300 × 100	0.005	240	240	27
	2	450 × 300 × 100	0.006	390	240	40
	3	750 × 500 × 130	0.009	630	420	146
	4	1500 × 1000 × 200	0.016	1100	700	900
	5	3000 × 2000 × 500	0.027	2000	1500	9000

Glass Surface Plates

Glass is an alternative material for surface plates. It was used during World War II when material and manufacturing capacity were in short supply. Glass can be ground suitably and has the benefit that it chips rather than raising a burr, which is a problem in cast iron surface plates.

4.4 V-BLOCKS

V-blocks are extensively used for inspection of jobs with a circular cross section. The major purpose of a V-block is to hold cylindrical workpieces to enable measurement. The cylindrical surface rests firmly on the sides of the 'V', and the axis of the job will be parallel to both the base and the sides of the V-block. Generally, the angle of the V is 90°, though an angle of 120° is preferred in some cases. It is made of high-grade steel, hardened above 60 Rc, and ground to a high degree of precision. V-blocks are manufactured in various sizes ranging from 50 to 200 mm. The accuracy of flatness, squareness, and parallelism is within 0.005 mm for V-blocks of up to 150 mm length, and 0.01 mm for those of length between 150 and 200 mm (Fig. 4.4).

V-blocks are classified into two grades, grade A and grade B, according to IS: 2949-1964, based on accuracy. Grade A V-blocks have minimum departure from flatness (up to 5 μm for 150 mm length) compared to grade B V-blocks.

There are many variants of V-blocks, such as V-blocks with clamp, magnetic V-block, and cast iron V-block. Figure 4.5 illustrates a V-block with a stirrup clamp. It is convenient for clamping the job onto the V-block, so that measurements can be made accurately. Another popular type of V-block is the magnetic V-block, shown in Fig. 4.6. The magnetic base sits on a flat surface, preferably on a surface plate. The base and two sides are energized for gripping onto a flat surface and a 'vee'slot enables the device to grip the job firmly with a circular cross section. A push-button control turns the permanent magnetic field on and off, thereby enabling the attachment or detachment of the V-block to a flat surface. All three magnetic surfaces are carefully ground and, when switched on, all three magnetic surfaces are activated simultaneously. Magnetic V-blocks are used in tool rooms for drilling and grinding round jobs.

4.5 GRADUATED SCALES

We often use the words 'rule' and 'scale' to mean the simple devices that we have been using since primary-school geometry class. However, there is a clear difference in the actual meaning of these two familiar words. A scale is graduated in *proportion* to a unit of length. For example,

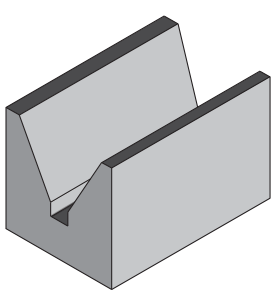


Fig. 4.4 V-block

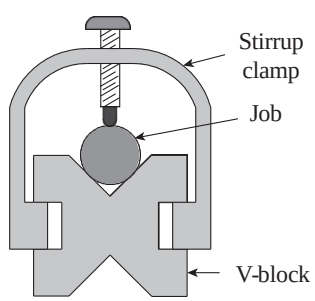


Fig. 4.5 V-block with a stirrup clamp

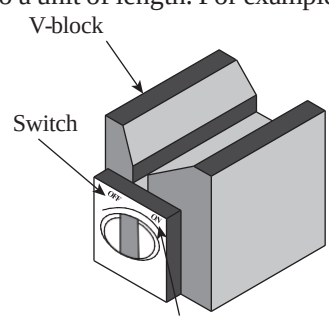


Fig. 4.6 Magnetic V-block

the divisions in an architect's scale, illustrated in Fig. 4.7, represent feet and inches, while the plumber's scale would have divisions in terms of 1/32th or 1/64th of an inch. The divisions of a rule, on the other hand, are the unit of length, its divisions, and its multiples. Typically, the rules with which we are familiar have graduations (in centimetres, millimetres, or inches) and their decimal divisions throughout the length.

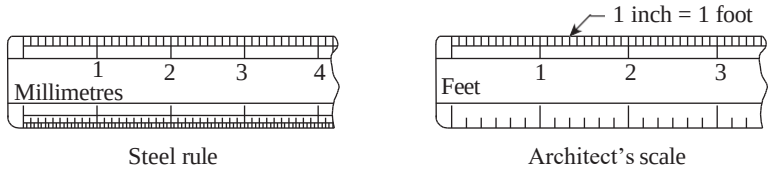


Fig. 4.7 Illustration of the difference between a rule and a scale

Steel rules are most popular in metrology applications since they are more accurate and durable compared to rules made from other materials such as wood or plastic. While rigid rules can be used for laying out lines on a job, flexible steel rules can also be used to measure surfaces with circular profiles. Steel rules are either stamped or cut from a roll of spring steel. The graduations are photo-engraved and tempered with satin chrome finish for good readability. The ruler can be 150, 300, 500, or 1000 mm long; 19 or 25 mm wide; and 1. mm thick. The finer sub-divisions may be marked either throughout the length of the scale or in only a part of its length.

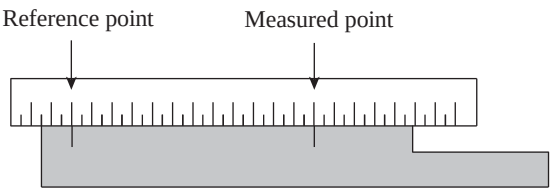


Fig. 4.8 Illustration of reference and measured points

The use of steel rule requires consideration of the relationship between the reference point and the measured point. Figure 4.8 illustrates the preferred way of choosing the reference point for making a measurement. A graduated line on the rule, rather than an edge of the rule, is selected as the reference point. This method improves the accuracy of measurement considerably, even though a little effort is required to align carefully, the reference and measured points. It is recommended not to use the edge of the rule as the reference point, as the edge is subjected to wear and tear and worn-out corners may contribute to error in measurements. Sometimes an attachment such as a hook or a knee is used to facilitate measurement, as shown in Fig. 4.9.

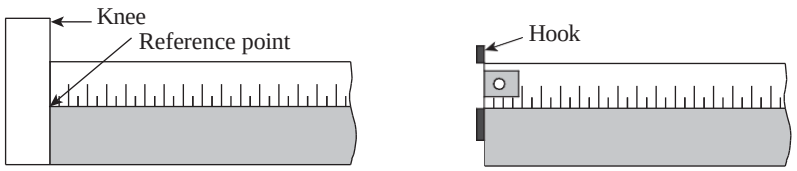


Fig 4.9 Correct ways of using a rule with attachments

4.5.1 Errors in Measurements

Steel rules are often used on the manufacturing shop floors for making measurements. Even though the least count of rules is limited to 1 mm or at the most to 0.5 mm, the user should

never be lax in the usage of any measuring device, however simple or trivial it may be. Therefore, one should be aware of the two common errors that can creep up in measurements involving rules. The first one is an inherent error because of poor quality of the rule. This can be attributed to either poor quality of the material of the rule or poor workmanship in the manufacture of the rule. One can avoid this by purchasing a good-quality rule from a standard vendor.

The second error may creep up because of wrong usage of a rule and the observational error. The rule should be properly aligned with the job being measured, ensuring that the reference and measured points are set accurately. The rule should not be twisted or bent in the process of measurement. The major observational error occurs because of parallax error. This is illustrated in Fig. 4.10. Parallax is the apparent shift in the position of an object caused by the change of position of the observer. If an observer views the scale along the direction B or C, the line of sight would be such that there is an apparent shift in the recorded reading by a division or two, as apparent from Fig. 4.10. The more the shift of the eye, from a vertical position right above the measured point, the more pronounced the error. It is needless to say that parallax error can be avoided if the observer recognizes this typical error and takes care to align his/her eyesight in the direction A, shown in Fig. 4.10.

Steel rules come in various sizes and shapes, depending upon the requirements of the component being measured. Accordingly, there are narrow rules, flexible fillet rules, short rules with holders, angle rules, measuring tapes, pocket tapes, and so on (Fig. 4.11). Narrow rules, fillet rules, and angle rules are used for measuring the inside of small holes, narrow slots, and grooves. Short rules with holders are convenient for measuring within recesses. Short rule is an extension of steel rule and obeys the same principle For very precise measurements, temperature expansion and contraction must be considered.

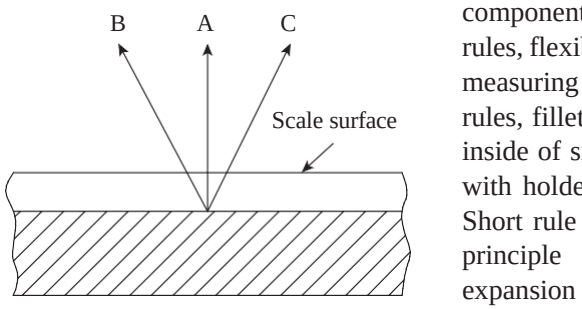


Fig. 4.10 Parallax error that can be minimized by direct eye measurements

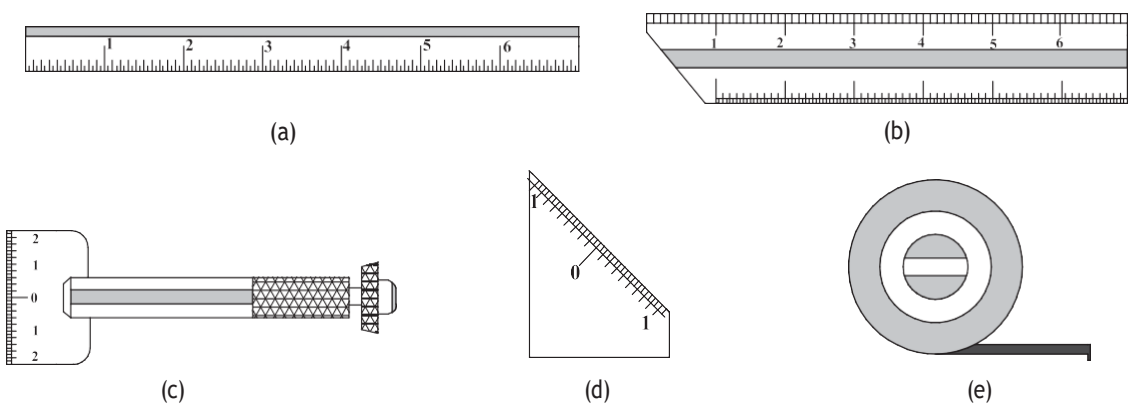


Fig. 4.11 Types of steel rules (a) Narrow tempered steel rule (b) Flexible fillet rule (c) Short rule with holder (d) Angle rule (e) Steel measuring tape

4.6 SCALED INSTRUMENTS

Rules are useful for many shop floor measurements. However, measurements of certain components require some mechanical means to either hold the measuring device steadily against the component being measured or capture the reading, which can be read at leisure. Another important advantage of a scaled instrument is that the least count of measurement can be improved greatly compared to an ordinary steel rule. Most of the modern scaled instruments provide digital display, which comes with a high degree of magnification. Measurements can be made up to micron accuracy. This section presents three scaled instruments, namely depth gauge, combination set, and callipers, which are necessary accessories in a modern metrology laboratory.

4.6.1 Depth Gauge

Depth gauge is the preferred instrument for measuring holes, grooves, and recesses. It basically consists of a graduated rod or rule, which can slide in a T-head (simply called the head) or stock. The rod or rule can be locked into position by operating a screw clamp, which facilitates accurate reading of the scale. Figure 4.12 illustrates a depth gauge, which has a graduated rule to read the measurement directly. The head is used to span the shoulder of a recess, thereby providing the reference point for measurement. The rod or rule is pushed into the recess until it bottoms. The screw clamp helps in locking the rod or rule in the head. The depth gauge is then withdrawn, and reading is recorded in a more convenient position. Thus, depth gauge is useful for measuring inaccessible points in a simple and convenient manner.

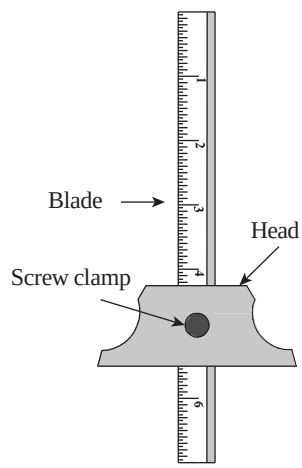


Fig. 4.12 Depth gauge

As already pointed out, either rods or rules can be used in depth gauges for the purpose of measurement. Although a slender rod can easily transfer measurements from narrow and inaccessible holes and recesses, the instrument cannot directly display the reading. One has to use another rule to measure the length of the protruded rod and record the measurement. This may lead to errors in measurements and reduce the reliability of the instrument. To overcome this problem, a graduated rod can be used, which can indicate the measurement directly. However, it is somewhat difficult to read graduations from a slender rod. Therefore, a narrow flat scale is the preferred choice for depth gauges. The rule is often referred to as the blade and is usually 150 mm long. The blade can accurately read up to 1 or ½ mm.

As already pointed out, the head is used to span the shoulder of a recess, thereby providing the reference point for measurement. This is illustrated in the rod-type depth gauge shown in Fig. 4.13. The end of the rod butts against the end surface to provide the measured point. Whenever depth needs to be measured, the projected length of the rod from the head is made very less. The lower surface of the head is firmly held against the job to ensure accurate location of the measured point. Now the rod is lowered until it butts against the surface of the job, thereby marking the measured point. The screw clamp is tightened, the instrument is slowly taken out, and the depth of the hole is read in a convenient position. This method is preferred

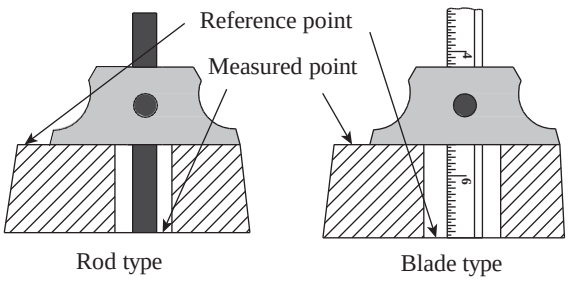


Fig. 4.13 Measured and reference points in depth gauges

for narrow recesses and holes. To summarize, the depth gauge is first positioned against the reference point, followed by the capture of the measured point in order to complete the measurement process.

Sometimes, it becomes necessary to alter the reference and measured points to suit the requirement, as illustrated by the blade-type depth gauge in Fig. 4.13. If the hole is large enough for visually positioning the blade of the depth gauge, the preferred method is to first locate the end of the blade against the lower surface of the hole. The blade is extended from the head, the instrument is brought close to the job, and the end of the blade is butted against the lower surface of the hole. This establishes the reference point for measurement. Now, the head is lowered and the lower surface of the head is made to butt against the top of the job, as shown in Fig. 4.10. The surface of the head provides the measured point. The screw clamp is now tightened and the measurement recorded.

Although depth gauge provides an easy and convenient method for measuring depths of holes and recesses, it has the following limitations:

1. The job size is limited by the width of the head of the depth gauge. Usually, the maximum width of the hole that can be spanned is about 50 mm.
2. The base of the head should be perpendicular to the line of measurement. Otherwise, the line of measurement will be skewed, resulting in erroneous readings.
3. The end of the blade must butt against the desired reference. This will be rather difficult to achieve, especially in blind holes.
4. The end of the blade and the lower surface of the head are always in contact with the job being measured. Therefore, these surfaces will undergo wear and tear. The instrument should be periodically checked for accuracy and replaced if the wear amounts to one graduation line of the instrument.

4.6.2 Combination Set

A combination set has three devices built into it: a combination square comprising a square head and a steel rule, a protractor head, and a centre head. While the combination square can be used as a depth or height gauge, the protractor head can measure the angles of jobs. The centre head comes in handy for measuring diameters of jobs having a circular cross section. The combination set is a useful extension of steel rule. This non-precision instrument is rarely used in any kind of production inspection. However, it is frequently used in tool rooms for tool and die making, pattern making, and fabrication of prototypes. It is a versatile and interesting instrument that has evolved from a try-square, which is used for checking squareness between two surfaces.

The graduated steel rule is grooved all along its length. The groove enables the square head to be moved along the length of the rule and fixed at a position by tightening the clamp screw provided on the square head. The square head along with the rule can be used for measuring heights and depths, as well as inside and outside squaring operations. The blade of the graduated

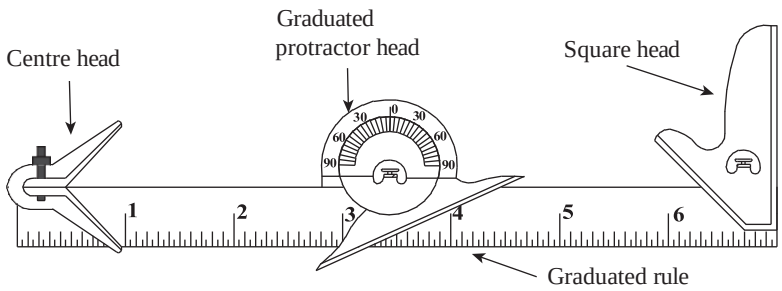


Fig. 4.14 Combination set

protractor head can be swivelled to any angle, which enables the measurement of angles on jobs. The protractor can also be moved along the scale and fixed at a convenient point. Protractors of some combination sets are provided with a spirit level for the purpose of levelling a surface. The centre head attachment is used with the rule to locate the centre of bar stocks. The illustration in Fig. 4.14 shows how each of these attachments are integrated in the combination set.

Square Head

The square head along with the graduated rule on the combination set provides an easy way of measuring heights and depths. While the square head provides a right angle reference, the rule provides a means for directly taking the readings. However, a primary requirement is that the square head can be used only against a flat reference surface. Figure 4.15 illustrates a typical method of measuring height using the combination set. The square head is firmly held against a flat surface of the job, and the rule is lowered until it touches the reference point at the bottom of the job, as shown in the figure. The rule can be locked in this position, and the reading noted down in a convenient position. Attachments are available to mark the measured point with reference to the end of the steel rule. The range of measurement can also be extended by using attachments. In some instruments, the square head is provided with a spirit level, which can be used to test the surfaces for parallelism. A scribing point is provided at the rear of the base in some instruments for scribing purposes.

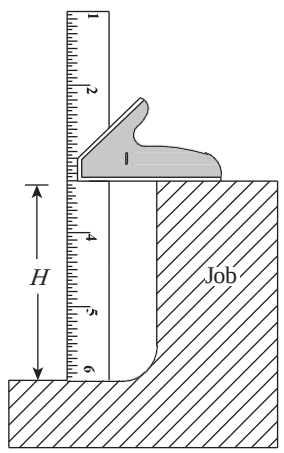


Fig. 4.15 Combination square as a height gauge

Protractor Head

Figure 4.16 illustrates the use of the protractor head on the combination set. This head comprises a rotatable turret within a stock. The turret has an angular scale graduated in degrees. Similar to the square head, the protractor head can also slide along the rule. The blade of the protractor is held firmly against the job and the angle can be directly read from the scale. A spirit level provided on the protractor head can be used for the purpose of levelling a surface. The protractor can also be used to determine the deviation of angle on the job from the desired one. The protractor is first set to the correct angle and locked in position. Now it is held against

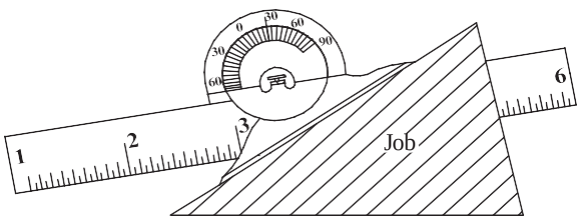


Fig. 4.16 Use of a protractor head for angle measurement

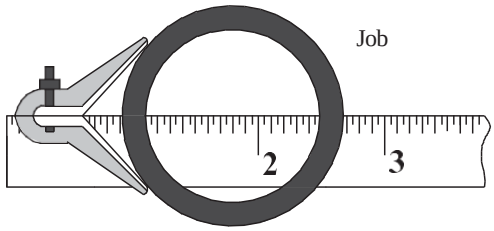


Fig. 4.17 Use of a centre head for the measurement of diameter

the surface of the job for which angle is to be measured. Any deviation from the desired angle can be checked by inserting angle gauges (feeler gauges) in the gap between the blade of the protractor and the job.

Centre Head

The centre head attachment is used along with the steel rule to locate the centre of a bar stock or a circular job. It can be seen from Fig. 4.17 that one edge of the steel rule bisects the V-angle of the centre head. Therefore, it lies on the centre line of any circular job held against the centre head. The diameter of the job can be directly read on the graduated scale, which is useful for marking the centre of the job by using a scriber. The V between the two blades of the centre head facilitates accurate positioning of circular jobs, which greatly improves measurement accuracy when compared to taking readings directly using a rule held against the job. The latter method is completely manual and depends on the skill of the person taking the reading, and therefore is highly prone to error.

Combination sets are available in various sizes. The length of the steel rule ranges from 150 to 600 mm. The scales are graduated in mm and 0.5 mm, and the graduations are available on both sides of the scale. Considering two sets of graduations on one side of the rule, it is possible to have four scales with different graduation schemes and least count. This enhances the versatility of the instrument considerably.

4.6.3 Callipers

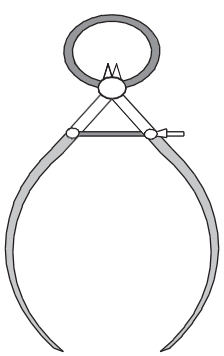


Fig. 4.18 Callipers, the original transfer instruments

There are many jobs whose dimensions cannot be accurately measured with a steel rule alone. A typical case in point is a job with a circular cross section. An attempt to take measurement using a steel rule alone will lead to error, since the steel rule cannot be positioned diametrically across the job with the required degree of accuracy. One option is to use the combination set. However, callipers are the original transfer instruments to transfer such measurements on to a rule (Fig. 4.18). They can easily capture the diameter of a job, which can be manually identified as the maximum distance between the legs of the calliper that can just slide over the diameter of the job. Even though callipers are rarely used in production inspection, they are widely used in tool room and related work.

Busch defines *callipers as instruments that physically duplicate the separation between the reference point and measured point of any dimension*

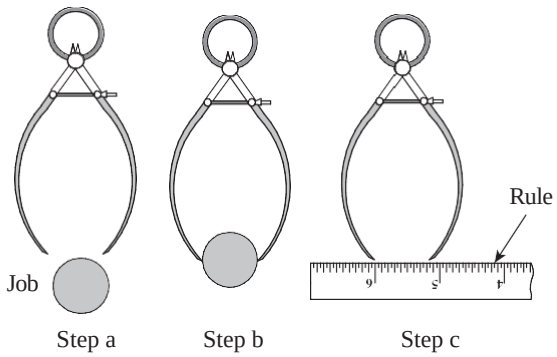


Fig. 4.19 Calliper being used to transfer a dimension from a job to a rule

within their range. Thus, callipers do only the job of transferring a dimension, but not of measuring instruments on their own. This is illustrated in Fig. 4.19, where a calliper is shown transferring the outer diameter of a job on to a graduated steel rule, to read the dimension accurately and conveniently. The outer diameter of a job is to be measured (Step a). Aligning the ends of the two legs of the calliper to a feature of the part being measured, like the one shown in Fig. 4.19, is accomplished quite easily (Step b) because the calliper provides for easy flexing of the two legs and a means of locking them into position whenever required. Now, simply laying the ends of the calliper on a steel rule facilitates easy measurement of the dimension in question (Step c). Thus, as the definition stated earlier mentions, physical duplication of the separation of reference and measured points is accomplished with a high degree of accuracy.

Callipers are available in various types and sizes. The two major types are the *firm joint calliper* and the *spring calliper*. A firm joint calliper, as the name itself suggests, can hold the position of two legs opened out to a particular degree unless moved by a certain force. This is possible because of higher friction developed at the joint between the two legs of the calliper. They are adjusted closely for size by gentle tapping of a leg. A locknut is needed to lock the calliper in a particular position. On the other hand, a spring calliper can hold a particular position due to the spring pressure acting against an adjusting nut. This permits a very careful control, and no lock is needed. Figure 4.20 illustrates the classification of callipers. Callipers are manufactured in a large number of sizes. They are designated not by their measurement ranges, but by the length of their legs,which range from 50 to 500 mm.

Figure 4.21 illustrates the different types of callipers. These are all simple callipers, whose ends are adjustable to transfer a measurement from the job to a steel rule. Although a member of the calliper family, a divider classified under callipers is simply referred to as a divider.

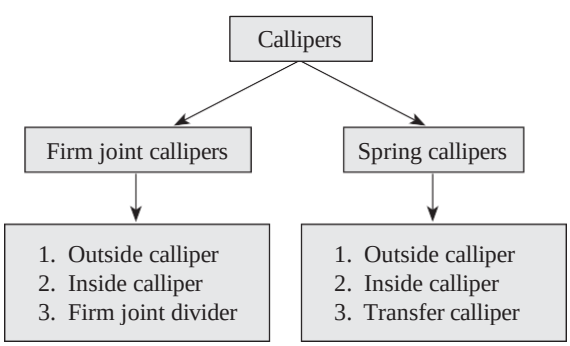


Fig. 4.20 Classification of callipers

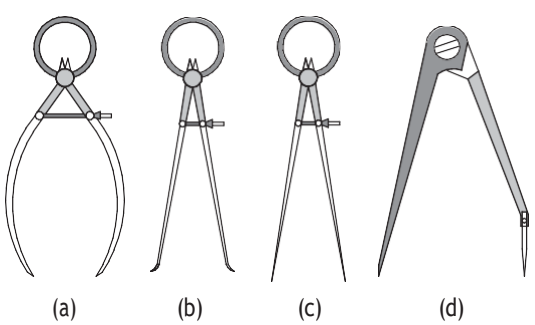


Fig. 4.21 Types of callipers (a) Outside calliper (b) Inside calliper (c) Divider (d) Hermaphrodite calliper

Outside calliper As the name suggests, an outside calliper is used to measure the outside or external dimensions. While taking measurements, the legs of the calliper should be set gently against the job either by a small physical force in case of a firm joint calliper or by an adjustment of the nut in case of a spring calliper. The user should ensure that a firm contact is established between the calliper and the job before transferring the measurement to a rule. One leg of the calliper is firmly held against a graduation on the steel rule to establish the reference point. The other leg is now gently transferred on to the rule to capture the measured point. The legs of callipers are made of alloy steel, with the measuring points being heat treated to withstand wear and tear. The legs are of a rectangular cross section, and should be free from cracks and any other type of flaw for longevity. The springs in spring callipers are made of spring steel, which is hardened and tempered. The spring force is adjusted by a knurled nut operating on a precision machined screw.

Inside calliper As illustrated in Fig. 4.21, the measuring ends of the legs of an inside calliper are shaped in the form of a hook. While taking measurements, the calliper legs are initially in a folded condition, which are inserted into the component. Now, the legs are stretched out to butt against the surface of the job. The calliper is carefully rocked over the centre. The feel provided as they pass the centre is the limit of their sensitivity.

Divider As with callipers, the primary use of dividers is to transfer measurements. However, they are used with line references rather than with surface references. Dividers are used for the following tasks: (a) transferring a dimension of a job to a rule for measurement, (b) transferring a dimension from one part to another part, and (c) transferring a dimension from a rule to a job for layout of the part. Scribing arcs during layout work is another chief use of dividers.

Hermaphrodite calliper It is essentially a scribing tool comprising one divider leg and one calliper leg. The scriber can be mounted by means of a locknut, as shown in Fig. 4.21. The chief advantage of a hermaphrodite calliper is that a scriber of any required shape and size can be fitted to it and used.

The proper use of the inside and outside callipers depends to a large extent on the skill of the person taking measurements. Measuring with a calliper consists of adjusting the opening so that its reference points duplicate the features of the job being measured. In other words, there is no other provision in a calliper that helps in its alignment than the reference points. As illustrated in Fig. 4.22, the greatest accuracy is achieved in case of callipers when the line of measurement coincides with a plane perpendicular to the job. The divider provides the best accuracy when the measurements are taken from well-marked lines, as shown in Fig. 4.22. Many a time measurements need to be taken between edges, in which case care must be exercised in ascertaining the proper way of taking measurements.

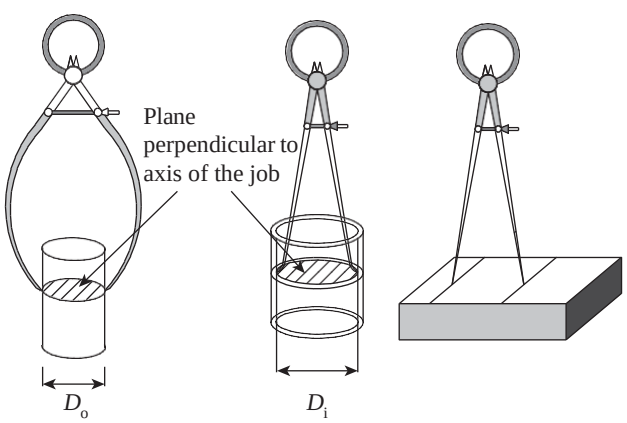


Fig. 4.22 Measurement using callipers (a) Outside calliper (b) Inside calliper (c) Divider

4.7 VERNIER INSTRUMENTS

The instruments discussed in this chapter until now can be branded 'non-precision' instruments, not for their lack of precision but for their lack of amplification. A steel rule can measure accurately up to 1 mm or at best up to 0.5 mm. It is not sensitive to variations in dimensions at much finer levels because of the inherent limitation in its design. On the other hand, vernier instruments based on the vernier scale principle can measure up to a much finer degree of accuracy. In other words, they can amplify finer variations in dimensions and can be branded as 'precision' instruments.

The vernier scale was invented in its modern form in 1631 by the French mathematician Pierre Vernier (1580–1637). Vernier instruments are being used for more than two centuries. The American, Joseph Brown, is credited with the invention of the vernier calliper. As is perhaps known to a student, a vernier scale provides a least count of up to 0.01 mm or less, which remarkably improves the measurement accuracy of an instrument. It has become quite common in the modern industry to specify dimensional accuracy up to 1 μm or less. It is the responsibility of an engineer to design and develop measuring instruments that can accurately measure up to such levels.

It will not be out of place here to briefly brush up our memory of the basic principles of a vernier scale. A vernier scale comprises two scales: the main scale and the vernier scale. Consider the scale shown in Fig. 4.23. Let us say that the main scale has graduations in millimetres up to a minimum division of 1 mm. The vernier scale also has graduations, having 10 equal divisions. In this example, 10 vernier scale divisions (VSDs) equal nine main scale divisions (MSDs). Obviously, the value of one VSD is less than one MSD. Such a vernier scale is called a *forward vernier*. On the other hand, suppose 10 VSDs equal 11 MSDs, the value of one VSD is more than that of one MSD. Such a vernier scale is called the *backward vernier*.

Calculation of least count The minimum length or thickness that can be measured with a vernier scale is called the *least count*. For a forward vernier shown in Fig. 4.23,

N VSD = $(N-1)$ MSD

1 VSD = $(N-1)/N$ MSD

Least count = 1 MSD − 1 VSD

Therefore, Least count = 1 MSD − $(N-1)/N$ MSD

Least count = $[1-(N-1)/N]$ MSD

Least count = 1 MSD/N

Total reading = MSR + (VC × LC), where MSR is the main scale reading, LC is the least count, and VC is the vernier coinciding division. Refer to Fig. 4.24 where the fourth division of the vernier coincides with a division on the main scale.

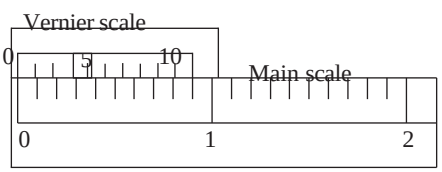


Fig. 4.23 Principle of a vernier scale

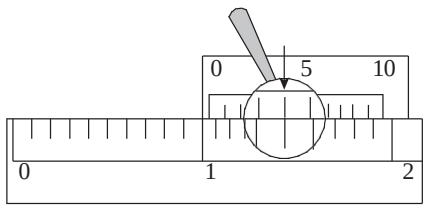


Fig. 4.24 Fourth division of vernier coinciding with a division on the main scale

Least count = 1 MSD/N = 1 mm/10 = 0.1 mm
Therefore, total reading = 1 + (4 × 0.1) = 1.4 mm

Digital read-out instruments and dial callipers are rapidly replacing vernier instruments. However, the principle of vernier measurement is basic to metrology, and the use of vernier instruments comprising vernier calliper, vernier depth gauge, vernier height gauge, vernier micrometers, etc., is still widespread in the industry. One can argue that anyone who can measure reliably with the vernier instruments can use the digital versions with equal reliability without any additional training. Even though the use of vernier instruments is not frequent for production inspection, they continue to play an important role in tool room and laboratory work. Production inspectors prefer limit gauges and comparators, which can speed up the inspection process considerably.

4.7.1 Vernier Calliper

A vernier calliper consists of two main parts: the main scale engraved on a solid L-shaped frame and the vernier scale that can slide along the main scale. The sliding nature of the vernier has given it another name—*sliding calliper*. The main scale is graduated in millimetres, up to a least count of 1 mm. The vernier also has engraved graduations, which is either a forward vernier or a backward vernier. The vernier calliper is made of either stainless steel or tool steel, depending on the nature and severity of application.

Figure 4.25 illustrates the main parts of a vernier calliper. The L-shaped main frame also serves as the fixed jaw at its end. The movable jaw, which also has a vernier scale plate, can slide over the entire length of the main scale, which is engraved on the main frame or the beam. A clamping screw enables clamping of the movable jaw in a particular position after the jaws have been set accurately over the job being measured. This arrests further motion of the movable jaw, so that the operator can note down the reading in a convenient position. In order to capture a dimension, the operator has to open out the two jaws, hold the instrument over the job, and slide the movable jaw inwards, until the two jaws are in firm contact with the job. A fine adjustment screw enables the operator to accurately enclose the portion of the job where measurement is required by applying optimum clamping pressure. In the absence of the fine adjustment screw, the operator has to rely on his careful judgement to apply the minimum force that is required to close the two jaws firmly over the job. This is easier said than done, since any excessive application of pressure increases wear and tear of the instrument and may also cause damage to delicate or fragile jobs. The two jaws are shaped in such a manner that

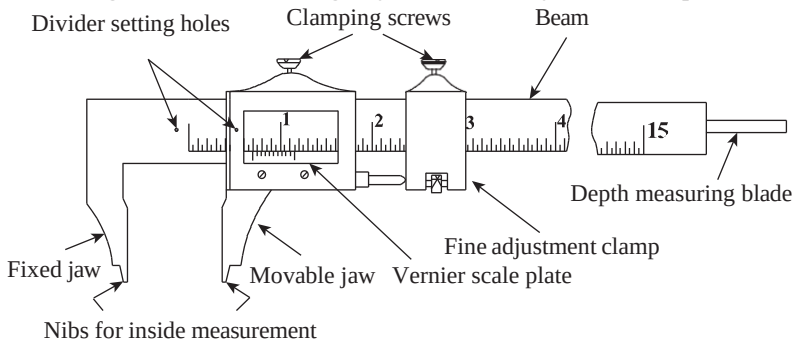


Fig. 4.25 Main parts of a vernier calliper

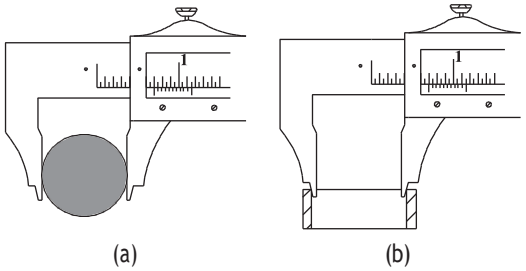


Fig. 4.26 Measurement of dimensions (a) Outside dimension (b) Inside dimension

they can be used to measure both inside and outside dimensions. Notice the nibs in Fig. 4.25, which can be used to measure inside dimension. Figure 4.26 illustrates the method of measuring inside and outside dimensions using a vernier calliper. Whenever the vernier slides over the main frame, a depth-measuring blade also slides in and out of the beam of the calliper. This is a useful attachment for measuring depths to a high degree of accuracy. Divider setting holes are provided, which enable the use of a divider to aid the measurement process.

Measuring a diameter is easier than measuring between flat surfaces, because the diameter is the greatest distance separating the reference and the measured points. Compared to the measurement between flat surfaces, the area of contact between the calliper and the job is much lesser in diameter measurement. Therefore, the resultant force acting either on the job or on the jaws of the calliper is lesser, with the result that there is no deformation or buckling of the jaws. This not only improves the accuracy of measurement, but also reduces the wear and tear of the instrument. Whether the measurement is done for the inside diameter or outside diameter, the operator has to rely on his/her feel to judge if proper contact is made between the measured surfaces and also that excessive force is not exerted on the instrument or the job. Continued closing of the calliper will increase the springing. High gauging pressure causes rapid wear of the jaws, burnishes the part (localized hardening of metal), and may cause damage to the calliper. The following guidelines are useful for the proper use of a vernier calliper:

1. Clean the vernier calliper and the job being measured thoroughly. Ensure that there are no burrs attached to the job, which could have resulted from a previous machining operation.
2. When a calliper's jaws are fully closed, it should indicate zero. If it does not, it must be recalibrated or repaired.
3. Loosen the clamping screw and slide the movable jaw until the opening between the jaws is slightly more than the feature to be measured.
4. Place the fixed jaw in contact with the reference point of the feature being measured and align the beam of the calliper approximately with the line of measurement.
5. Slide the movable jaw closer to the feature and operate the fine adjustment screw to establish a light contact between the jaws and the job.
6. Tighten the clamp screw on the movable jaw without disturbing the light contact between the calliper and the job.
7. Remove the calliper and note down the reading in a comfortable position, holding the graduations on the scale perpendicular to the line of sight.
8. Repeat the measurement a couple of times to ensure an accurate measurement.
9. After completing the reading, loosen the clamping screw, open out the jaws, and clean and lubricate them.
10. Always store the calliper in the instrument box provided by the supplier. Avoid keeping the vernier calliper in the open for long durations, since it may get damaged by other objects or contaminants.
11. Strictly adhere to the schedule of periodic calibration of the vernier calliper.

According to IS: 3651-1974,vernier callipers are of three types: type A, type B, and type C. While all the three types have the scale on the front of the beam, type A vernier scale has jaws on both sides for external and internal measurements, along with a blade for depth measurement. Type B, shown in Fig. 4.25, is provided with jaws on one side only for both external and internal measurements. Type C has jaws on both sides for making the measurements. However, the jaws have knife edge faces for marking purpose. The recommended measuring ranges for vernier callipers are 0–125, 0–200, 0–250, 0–300, 0–500, 0–750, 0–1000, 750–1500, and 750–2000 mm.

Dial Calliper

A vernier calliper is useful for accurate linear measurements. However, it demands basic mathematical skill on the part of the user. One should be able to do simple calculations involving MSD, vernier coinciding division, and least count, in order to compute the measured value of a dimension. In addition, considerable care should be exercised in identifying the coinciding vernier division. These problems can be offset by using a dial calliper (Fig. 4.27).

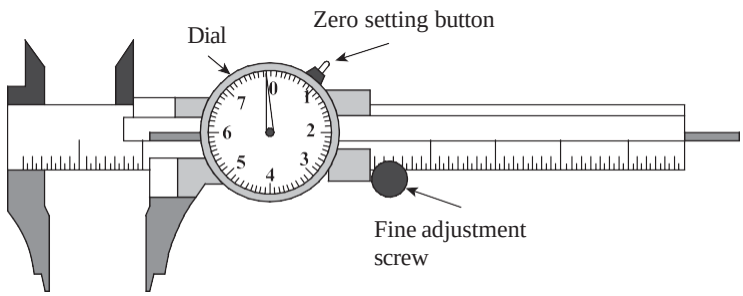


Fig. 4.27 Dial calliper

In a dial calliper, the reading can be directly taken from a dial gauge that is attached to the calliper. The dial gauge has its own least count, which is clearly indicated on the face of the dial. By multiplying the value of the reading indicated by the least count, one can calculate the measured value easily. A small but precise pair of rack and pinion drives a pointer on a circular scale. This facilitates direct reading without the need to read a vernier scale. Typically, the pointer undergoes one complete rotation per centimetre or per millimetre of linear measurement. This measurement should be added to the main scale reading to get the actual reading. A dial calliper also eliminates parallax error, which is associated with a conventional vernier calliper.

A dial calliper is more expensive than the vernier calliper. In addition, the accuracy of the reading mechanism of the dial calliper is a function of length of travel, unlike the vernier calliper that has the same accuracy throughout its length. A dial calliper is also subject to malfunctioning because of the delicate nature of the dial mechanism.

Electronic Digital Calliper

An electronic digital calliper is a battery-operated instrument that displays the reading on a liquid crystal display (LCD) screen. The digital display eliminates the need for calculations and provides an easier way of taking readings. Figure 4.28 illustrates the main parts of an electronic digital calliper.

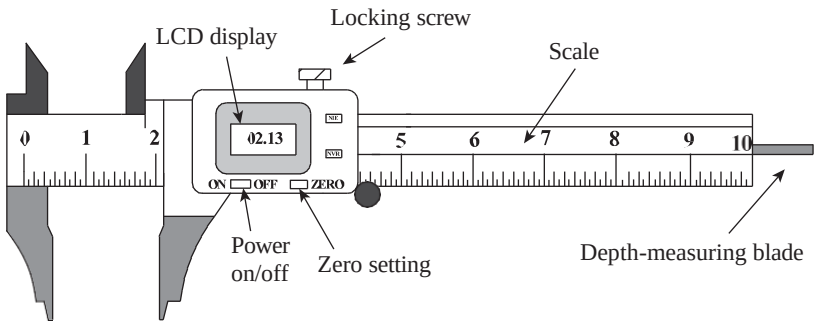


Fig. 4.28 Electronic digital calliper

The LCD display is turned on or off with a button. In order to initialize the instrument, the external jaws are brought together until they touch and the 'zero button' is pressed to set the reading to zero. The digital calliper can now be used to measure a linear dimension. Some digital callipers can be switched between centimetres or millimetres, and inches. Digital callipers are made of stainless steel and are generally available in three sizes: 150, 200, and 300 mm.

The two greatest advantages of an electronic digital calliper are its electronic calculator functions and capability to be interfaced with a computer. It can be set to either metric or British system of units. The 'floating zero' option allows any place within the scale range to be set to zero. The digital display will then exhibit either plus or minus deviations of the jaw from a reference value. This enables the instrument to be also used as a limit gauge. More importantly, a digital calliper can be interfaced with a dedicated recorder or personal computer through a serial data cable. The digital interface provides secured storage for a series of readings, thereby improving the reliability of the records. It can be connected to a printer to provide a printed record or can be directly interfaced with a computer of a statistical control system.

4.7.2 Vernier Depth Gauge

Section 4.7.1 already highlighted the construction and working principle of a depth gauge. However, even though it is simple to operate, it cannot take measurements finer than 1 mm accuracy. A vernier depth gauge is a more versatile instrument, which can measure up to 0.01 mm or even finer accuracy. Figure 4.29 illustrates the constructional features of a vernier depth gauge. The lower surface of the base has to butt firmly against the upper surface of the hole or recess whose depth is to be measured. The vernier scale is stationary and screwed onto the slide, whereas the main scale can slide up and down. The nut on the slide has to be loosened to move the main scale. The main scale is lowered into the hole or recess, which is being measured. One should avoid exerting force while pushing the scale against the surface of the job being measured, because this will not only result in the deformation of the scale resulting in erroneous measurements, but also accelerate the wear and tear of the instrument. This problem is eliminated thanks to the fine adjustment clamp provided with the instrument. A fine adjustment wheel will rotate the fine adjustment screw, which in turn will cause finer movement of the slide. This ensures firm but delicate contact with the surface of the job.

Vernier depth gauges can have an accuracy of up to 0.01 mm. Periodic cleaning and lubrication are mandatory, as the main scale and fine adjustment mechanism are always in motion in the process of taking measurements.

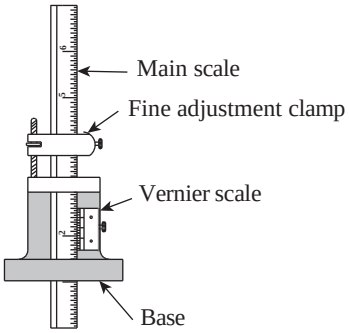


Fig. 4.29 Vernier depth gauge

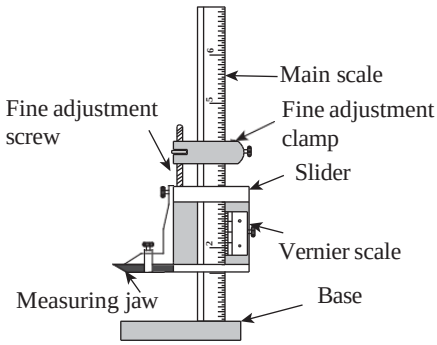


Fig. 4.30 Vernier height gauge

4.7.3 Vernier Height Gauge

In a vernier height gauge, as illustrated in Fig. 4.30, the graduated scale or bar is held in a vertical position by a finely ground and lapped base. A precision ground surface plate is mandatory while using a height gauge. The feature of the job to be measured is held between the base and the measuring jaw. The measuring jaw is mounted on a slider that moves up and down, but can be held in place by tightening of a nut. A fine adjustment clamp is provided to ensure very fine movement of the slide in order to make a delicate contact with the job. Unlike in depth gauge, the main scale in a height gauge is stationary while the slider moves up and down. The vernier scale mounted on the slider gives readings up to an accuracy of 0.01 mm.

Vernier height gauges are available in sizes ranging from 150 to 500 mm for precision tool room applications. Some models have quick adjustment screw release on the movable jaw, making it possible to directly move to any point within the approximate range, which can then be properly set using the fine adjustment mechanism. Vernier height gauges find applications in tool rooms and inspection departments. Modern variants of height gauges such as optical and electronic height gauges are also becoming increasingly popular.

4.8 MICROMETER INSTRUMENTS

The word ‘micrometer’ is known by two different meanings. The first is as a unit of measure, being one thousandth of a millimetre. The second meaning is a hand-held measuring instrument using a screw-based mechanism. The word *micrometer* is believed to have originated in Greece, the Greek meaning for this word being *small*. The first ever micrometer screw was invented by William Gascoigne of Yorkshire, England, in the 17th century and was used in telescopes to measure angular distances between stars. The commercial version of the micrometer was released by the Browne & Sharpe Company in the year 1867. Obviously, micrometer as an instrument has a long and cherished history in metrological applications. There have been many variants of the instrument, and modern industry makes use of highly sophisticated micrometers, such as digital micrometers and laser scan micrometers. A micrometer can provide better least counts and accuracy than a vernier calliper. Better accuracy results because of the fact that the line of measurement is in line with the axis of the instrument, unlike the vernier calliper that does not conform to this condition. This fact is best explained by *Abbe’s principle,* which states that *‘maximum accuracy may be obtained only when the standard is in line with the axis of the*

part being measured'. Figure 4.31 illustrates the relevance of Abbe's law for micrometers and vernier callipers.

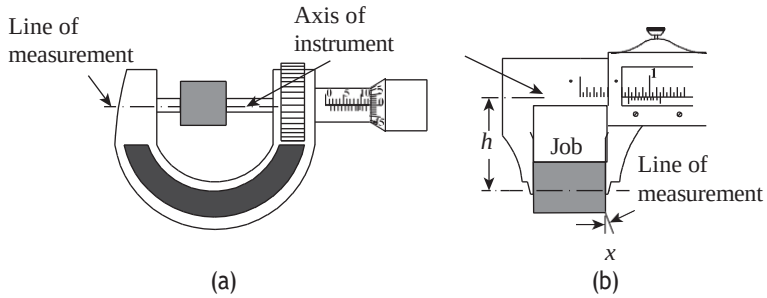


Fig. 4.31 Conformity to Abbe's law (a) Micrometer (b) Vernier calliper

In case of a micrometer, the axis of the job being measured is in line with the line of measurement of the instrument, as illustrated in Fig. 4.31(a). In case of a vernier calliper, for the reading to be accurate, the beam would have to be perfectly straight and the two jaws perfectly at 90° to it. However, this is rarely the case. There is always some lack of straightness of the beam, and the jaws may not be perfectly square with the beam. With continuous usage and wear and tear, the jaws will develop more and more play (Play refers to uncontrolled movements due to slip of one part over the other.) because of repeated sliding movements. Therefore, a certain amount of angular error, marked as *x* in Fig. 4.31(b), will always be present. This angular error also depends on how far the line of measurement is from the axis of the instrument. The higher the value of this separation *h*, the greater will be the angular error. We can therefore conclude that *the degree to which an instrument conforms to Abbe's law determines its inherent accuracy.*

4.8.1 Outside Micrometer

Figure 4.32 illustrates the details of an outside micrometer. It consists of a C-shaped frame with a stationary anvil and a movable spindle. The spindle movement is controlled by a precision ground screw. The spindle moves as it is rotated in a stationary spindle nut. A graduated scale is engraved on the stationary sleeve and the rotating thimble. The zeroth mark on the thimble will coincide with the zeroth division on the sleeve when the anvil and spindle faces are brought together. The movement of the screw conforms to the sets

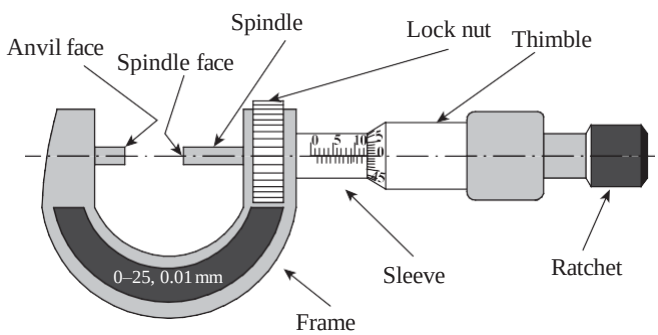


Fig. 4.32 Outside micrometer

of graduations. The locknut enables the locking of the spindle while taking a reading. The ratchet ensures a 'feel' while taking a reading and prevents application of excessive force on the job. The ranges of micrometers are normally 0–25, 25–50, or 0–50 mm. The maximum range of micrometers is limited to 500 mm.

A micrometer is made of steel or cast steel. The measuring faces are hardened to about 60–65 HRC since they are in constant touch with metallic jobs being measured. If warranted, the faces are also tipped with tungsten carbide or a similar material to prevent rapid wear. The anvil is ground and lapped to a high degree of accuracy. The material used for thimble and ratchet should be wear-resistant steel.

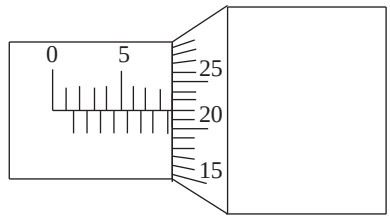


Fig. 4.33 Reading an outside micrometer

Micrometers with metric scales are prevalent in India. The graduations on the sleeve are in millimetres and can be referred to as the main scale. If the smallest division on this scale reads 0.5 mm, each revolution of the thimble advances the spindle face by 0.5 mm. The thimble, in turn, will have a number of divisions. Suppose the number of divisions on the thimble is 50, then the least count of the micrometer is 0.5/50, that is, 0.01 mm. Figure 4.33 illustrates how the micrometer scale is read when a job is held between the anvil face and the spindle face.

In this example, the main scale reading is 8.5 mm, which is the division immediately preceding the position of the thimble on the main scale. As already pointed out, let us assume the least count of the instrument to be 0.01 mm. The 22nd division on the thimble is coinciding with the reference line of the main scale. Therefore, the reading is as follows:

$$8.5 + 22\,(0.01)\,\text{mm} = 8.72\,\text{mm}$$

Thus, a micrometer is a simple instrument to use. However, there are two precautions to be observed while reading a micrometer. The thimble must be read in the correct direction. The other precaution concerns the zero position on the thimble. When passing the index line on the main scale, there is a chance to read an extra 0.5 mm. This is caused by the fact that the next main scale graduation has begun to show but has not yet fully appeared. This is avoided by being careful to read only full divisions on the barrel. Assuming that these simple precautions are adhered to, a micrometer has many advantages over other linear measurement instruments. It has better readability than a vernier scale and there is no parallax error. It is small, lightweight, and portable. It retains accuracy over a longer period than a vernier calliper and is less expensive. On the flip side, it has a shorter measuring range and can only be used for end measurement.

Types of Micrometers

A micrometer is a versatile measuring instrument and can be used for various applications by simply changing the anvil and the spindle face. For example, the anvil may be shaped in the form of a V-block or a large disk. Figure 4.34 shows a few variants, namely the disk micrometer, screw thread micrometer, dial micrometer, and blade micrometer. The following paragraphs briefly highlight the use of each type of micrometer in metrology applications:

Disk micrometer It is used for measuring the distance between two features with curvature. A tooth span micrometer is one such device that is used for measuring the span between the two teeth of a gear. Although it provides a convenient means for linear measurement, it is prone to error in measurement when the curvature of the feature does not closely match the curvature of the disk.

Screw thread micrometer It measures pitch diameters directly. The anvil has an internal '*vee*', which fits over the thread. Since the anvil is free to rotate, it can accommodate any rake range of thread. However, interchangeable anvils need to be used to cover a wide range of thread pitches. The spindle has a conical shape and is ground to a precise dimension.

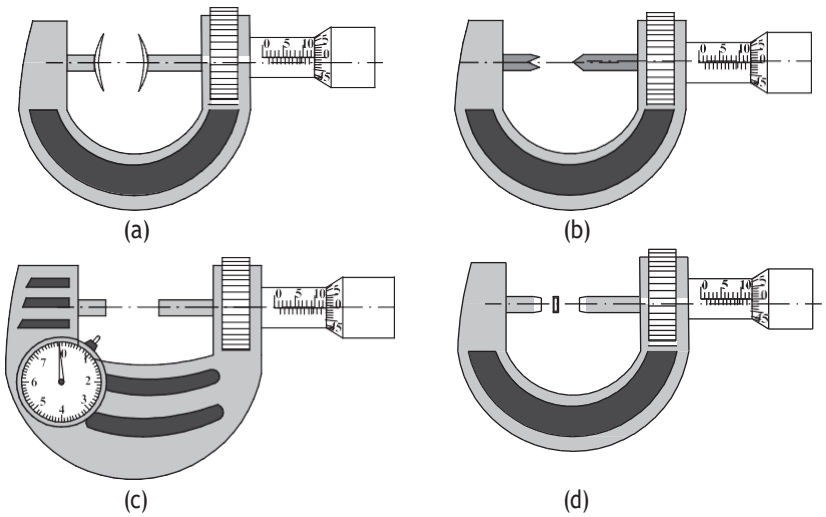


Fig. 4.34 Types of micrometers (a) Disk type (b) Screw thread type (c) Dial type (d) Blade type

Dial micrometer The dial indicator fixed to the frame indicates the linear displacement of a movable anvil with a high degree of precision. It is especially useful as a comparator for GO/NO-GO judgement in mass production. The dial micrometer normally has an accuracy of 1 μm and repeatability of 0.5 μm. Instruments are available up to 50 mm measuring distance, with a maximum measuring force of 10 N. The dial tip is provided with a carbide face for a longer life.

Blade micrometer The anvil and spindle faces are in the form of narrow blades and useful for measuring narrow grooves, slots, keyways, and recesses. The blade thickness is around 0.75–1 mm. The spindle does not rotate when the movable blade is moving along the measuring axis. Due to the slender nature of the instrument and non-turning spindle working against a rotating screw, it is vulnerable to rapid wear and tear and needs careful use and maintenance.

Universal micrometer It has interchangeable anvils such as flat, spherical, spline, disk, or knife edge. It is called universal because of its modular design. The micrometer fitted with the required accessories can function as an outside micrometer, a depth micrometer, a step micrometer, etc.

4.8.2 Vernier Micrometer

A micrometer that we considered hitherto can provide an accuracy of at best 0.01 mm or 10 μm. Placing a vernier scale on the micrometer permits us to take readings up to the next decimal place. In other words, one can accurately measure up to 1 μm or 0.001 mm, which is an excellent proposition for any precision workmanship. As illustrated in Fig. 4.35, in addition to the barrel and thimble scales, a vernier scale is provided next to the barrel scale. Divisions on this vernier scale have to be read in conjunction with the barrel scale to provide the next level of discrimination in readings. The vernier scale consists of a number of equally spaced lines, which are numbered from 0 to 5 or 10, depending on the scale.

The principle of measurement of a vernier micrometer is very similar to that of a vernier calliper. If a division on the thimble is exactly coinciding with the reference line (line marked 0 in the vernier scale in Fig. 4.35) of the vernier scale, the reading is taken in a way similar to an ordinary micrometer explained earlier. However, if none of the divisions on the thimble coincide with the reference line, we need to examine which division on the thimble coincides with one of the divisions on the vernier scale. Hence, an additional step is involved in the calculation since the vernier reading should be taken into account.

Refer to Fig. 4.36, which shows a sample reading. In this case, the thimble has crossed the 12.5 mm mark on the barrel scale. None of the divisions on the thimble coincides with the zeroth line on the vernier scale, that is, the reference line on the barrel. However, the reference line is between the 24th and 25th divisions on the thimble. Suppose the thimble has 50 divisions, and five divisions on the vernier scale correspond to six divisions on the thimble, we can calculate the least count of the instrument as follows.

If one complete rotation of the thimble moves it by 0.5 mm on the barrel scale, the least count of the micrometer scale is 0.5/50 = 0.01 mm.

Since five divisions on the vernier scale correspond to six divisions on the thimble, the least count of the vernier scale is equal to 0.01/5 = 0.002 mm. In Fig. 4.36, the fourth division on the vernier scale is coinciding with a division on the thimble.

Therefore, the reading is 12.5 + 24 (0.01) + 4 (0.002) = 12.748 mm.

Guidelines for Use of Micrometers

1. Before placing the micrometer on the job being measured, bring it near the desired opening. Do this by rolling the thimble along the hand but not by twirling. Hold the micrometer

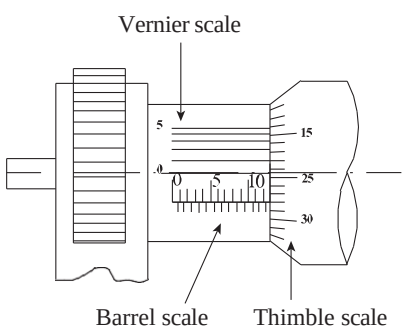


Fig. 4.35 Vernier micrometer

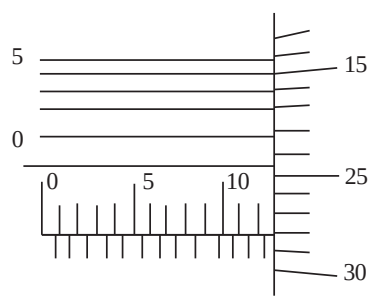


Fig. 4.36 Reading a vernier scale

firmly with one hand, and use the feel of the hand to ensure that the axis of the micrometer is perpendicular to the reference plane of the job. Close the micrometer using the ratchet stop until it disengages with a click.

2. Even though a micrometer can be locked in position by tightening the clamp ring (locknut) and used as a snap gauge for inspection purposes, it is not basically designed for this role. Locking the spindle movement and forcing the measuring faces over the job result in sliding friction, which accelerates wear on the contact surfaces as well as on the micrometer screw.
3. The locknut is a memory device. It retains the reading so that it can be read in a convenient position. However, avoid tightening the locknut when the spindle is withdrawn. Doing so will injure the clamping mechanism.
4. It is not wise to buy a micrometer that does not have a controlled force feature. Excessive force while closing the measuring faces over the job will result in rapid wear and tear of the instrument. A ratchet stop acts as an overriding clutch that holds the gauging force at the same amount for each measurement regardless of the differences in manual application of force.
5. While measuring the diameter of a cylindrical part, rock the cylinder to find the maximum opening that provides the desired feel.
6. Do not expect the micrometer to guarantee reliable measurement if it is (a) dirty; (b) poorly lubricated; (c) poorly adjusted; or (d) closed too rapidly.
7. At the end of each day, the micrometer should be wiped clean, visually inspected, oiled, and replaced in its case to await the next use.

4.8.3 Digital Micrometer

The 'multifunction' digital micrometer is becoming very popular in recent times. The readings may be processed with ease. The push of a button can convert a reading from decimal to inch and vice versa. Any position of the spindle can be set to zero and the instrument can be used to inspect a job within a specified tolerance. The instrument can be connected to a computer or a printer. Most instruments can record a series of data and calculate statistical information such as mean, standard deviation, and range (Fig. 4.37).

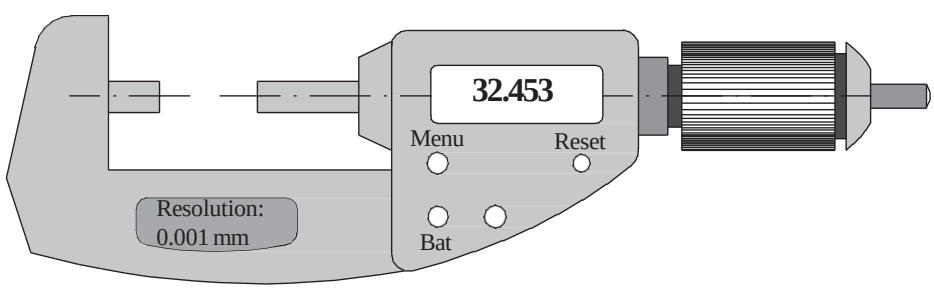


Fig. 4.37 Digital micrometer

The instrument is recommended to be used along with a stand for ease of measurement. The spindle is made of stainless steel and measuring faces are carbide tipped for a longer life. A locking clamp ensures locking of spindle at any desired setting. A constant and low measuring force is ensured by the thimble mechanism. Most of the instruments have a least count of

0.001 mm. An LCD screen displays the reading with absolute linear scale with simple digimatic data collection for personal computer (SPC) data output. An easy push button control is provided to choose the various functions of the instrument. The push buttons controlling the various functions are as follows:

1. ON/OFF: To power the instrument on or off
2. IN/MM: To select either inch or metric system of measurement
3. ZERO: To set the screen display to zero at any desired position
4. HOLD: To hold the measurement taken until the push button is operated again
5. ORIGIN: To set the minimum value for the micrometer depending upon its size; micrometer count commences from this value
6. Alarm indicator: To indicate low voltage and counting value composition error

4.8.4 Inside Micrometer Calliper

The inside micrometer calliper is useful for making small measurements from 5 to 25 mm. In this instrument, unlike a regular micrometer, the axis of the instrument does not coincide with the line of measurement. In addition, unlike the outside micrometer where there is a surface contact between the job and the instrument, the contact between the job and the instrument is line contact. The *nibs*, as the contacts are called, are ground to a small radius. As a necessity, this radius has to be smaller than the smallest radius the instrument can measure. Therefore, all measurements are made with line contacts.

As illustrated in Fig. 4.38, the movable jaw can be moved in and out by the rotation of the thimble. One complete rotation of the thimble moves it by one division on the barrel scale. A locknut can be operated to hold the position of the movable jaw for ease of noting down a reading. While taking measurements, it needs to be rocked and centralized to assure that the axis of the instrument is parallel to the line of measurement. This makes the instrument prone to rapid wear. It is therefore needless to say that the instrument needs to be checked and calibrated regularly.

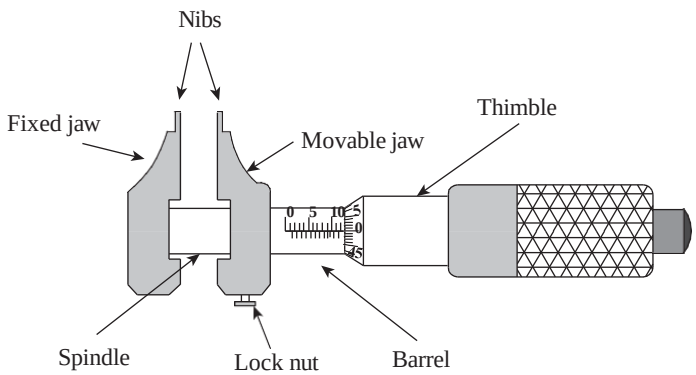


Fig. 4.38 Inside micrometer calliper

4.8.5 Inside Micrometer

This instrument perfectly complies with Abbe's law. The axis of an inside micrometer is also its line of measurement. It is useful for measuring the inside diameter of cylinders, rings, and other machine parts. The inside micrometer set has several accessories, which have to be assembled together for taking the readings. The main unit is the measuring head, which has a thimble that moves over a barrel, same as in the case of an outside micrometer. Graduated scales are provided on the barrel and thimble, which give readings up to an accuracy of 0.01 mm, but with a limited range. The rear end of the measuring head has a contact surface, whereas extension

rods of various lengths can be fitted to the front end of the measuring head. A set of extension rods are provided with the instrument to cover a wide range of measurements. The rod ends are spherical and present nearly point contact to the job being measured. A chuck attached to the spindle facilitates the attachment of extension rods. Using a particular extension rod, the distance between contact surfaces can be varied by rotating the thimble up to the range of the micrometer screw. Higher diameters and distances can be measured using longer extension rods. Figure 4.39 illustrates the construction details of an inside micrometer.

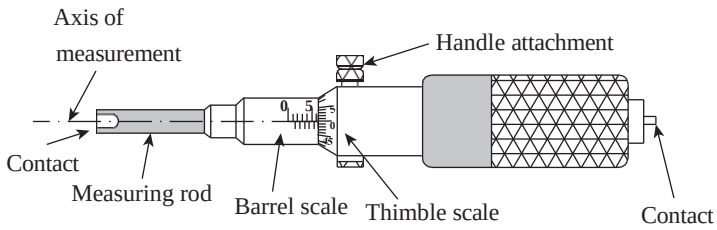


Fig. 4.39 Inside micrometer

The inside micrometer is more common than the inside micrometer calliper because of the flexibility it affords to measure a large range of dimensions. A range of 25 mm above the length of extension rods is commonly used in shops. A standard set will have five extension rods of ranges 50–75, 75–100, 100–125, 125–150, 150–175, and 175–200 mm. In addition to extension rods, a spacing collar (usually 12.5 mm in length) is provided for smaller adjustments in the range of measurements. The micrometer is also provided with a detachable handle for easier handling of the instrument.

The best practice for using an inside micrometer is to first measure the dimension approximately by using a steel rule. Then select a suitable extension rod that can adequately cover the range of measurements. Insert the extension rod into the chuck and set the instrument to read zero. Now fix the handle and lower the instrument into the gap where the dimension is to be measured. Operate the thimble until the two contact surfaces establish a firm contact with the surfaces of the job. While measuring diameters, it is always recommended to lightly move the instrument to and fro so that the actual diameter is sensed by the person taking measurements. Now, operate the locknut and take out the instrument. The micrometer reading has to be added to the length of the extension rod to get the actual reading.

4.8.6 Depth Micrometer

An alternative to vernier depth gauge is the depth micrometer. In fact, most shop floor engineers vouch for its superiority over vernier depth gauges because of its greater measuring range, better reliability, and easier usability. One peculiarity of this instrument is that it reads in reverse from other micrometers. Looking from the ratchet side, a clockwise rotation moves the spindle downwards, that is, into the depth of the job being measured. Therefore, the entire barrel scale is visible when the tip of the measuring rod is in line with the bottom surface of the base. As the measuring rod advances into the depths, the thimble will move over the barrel scale. Reliable measurements of up to 0.01 mm are possible with this instrument. Figure 4.40 illustrates the parts of a depth micrometer. The bottom flat surface of the base butts over the reference plane on the job, and the micrometer scale directly gives the depth of the measuring rod tip from the reference plane.

The head movement of the depth micrometer is usually 25 mm. Inter-changeable measuring rods, similar to an inside micrometer discussed in the previous section, provide the required measuring range for the instrument. Measuring rods of up to 250 mm length are used in a standard set.

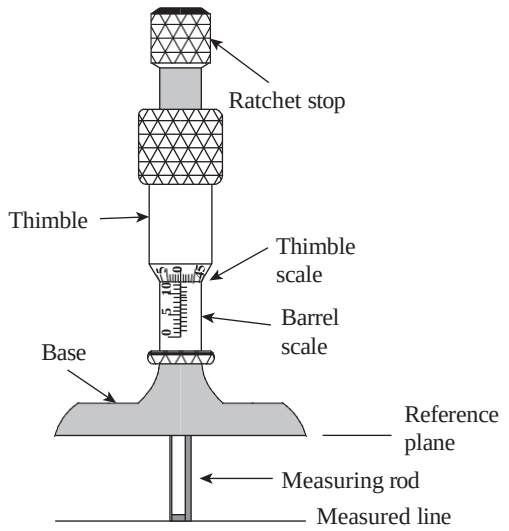


Fig. 4.40 Depth micrometer

4.8.7 Floating Carriage Micrometer

A floating carriage micrometer, sometimes referred to as an effective diameter-measuring micrometer, is an instrument that is used for accurate measurement of 'thread plug gauges'. Gauge dimensions such as outside diameter, pitch diameter, and root diameter are measured with the help of this instrument. All these dimensions have a vital role in thread plug gauges, since the accuracy and interchangeability of the component depend on the gauges used. To reduce the effect of slight errors in the micrometer screws and measuring faces, this micrometer is basically used as a comparator. Figure 4.41 illustrates a floating carriage micrometer.

The carriage has a micrometer with a fixed spindle on one side and a movable spindle with the micrometer on the other side. The carriage moves on a finely ground 'V' guide way or an antifriction guide way to facilitate movement in a direction parallel to the axis of the plug gauge mounted between the centres. The micrometer has a non-rotary spindle with a least count of up to 0.001 or 0.002 mm. The instrument is very useful for thread plug gauge manufacturers, in gauge calibration laboratories (established under NABL accreditation), and in standard rooms where in-house gauge calibration is carried out.

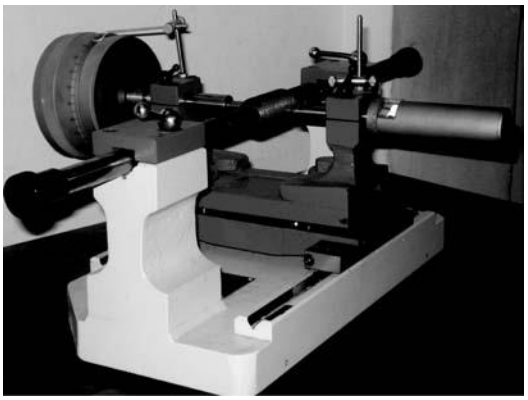


Fig. 4.41 Floating carriage micrometer

4.9 SLIP GAUGES

Hitherto we have seen instruments such as vernier calliper, depth gauge, and micrometer, which can facilitate measurement to a fairly high degree of accuracy and precision. All these measurements involve line standards. The accuracy of these instruments depends on the accuracy of the workmanship involved in their manufacture. Any minor misalignment or error in a screw can lead to errors in measurement. Repetitive use of a screw or joint results in rapid

wear and tear, which can lead to accumulation of errors in measurement within a short time. Slip gauges, also called gauge blocks, can counter some of these limitations and provide a high degree of accuracy as *end standards*. In fact, slip gauges are a direct link between the measurer and the international length standards.

The origin of gauge blocks can be traced to the 18th century Sweden, where 'gauge sticks' were known to have been used in machine shops. However, the modern-day slip gauges or gauge blocks owe their existence to the pioneering work done by C.E. Johansson, a Swedish armoury inspector. Therefore, gauge blocks are also known as *Johansson gauges*. He devised a set of slip gauges manufactured to specific heights with a very high degree of accuracy and surface finish. He also proposed the method of 'wringing' slip gauges to the required height to facilitate measurements. He also emphasized that the resulting slip gauges, to be of universal value, must be calibrated to the international standard. Johansson was granted a patent for his invention in the year 1901 and formed the Swedish company CE Johansson AB in the year 1917. He started manufacturing and marketing his gauge blocks to the industry, and found major success in distant America. One of his customers was Henry Ford with whom he signed a cooperative agreement to establish a gauge making shop at his Cadillac automobile company. The development of 'GO' and 'NO-GO' gauges also took place during this time.

Figure 4.42 illustrates the functional features of a slip gauge. It is made of hardened alloy steel having a 30 mm × 10 mm cross section. Steel is the preferred material since it is economical and has the same coefficient of thermal expansion as a majority of steel components used in production. Hardening is required to make the slip gauge resistant to wear. Hardening is followed by stabilizing at a sub-zero temperature to relieve stresses developed during heat treatment. This is followed by finishing the measuring faces to a high degree of accuracy, flatness, and surface finish. The height of a slip gauge is engraved on one of the rectangular faces, which also features a symbol to indicate the two measured planes. The length between the measuring surfaces, flatness, and surface conditions of measuring faces are the most important requirements of slip gauges.

Carbide gauge blocks are used for their superior wear resistance and longer life. They also have low coefficient of thermal expansion. However, they are quite expensive and used when rapid wear of gauges is to be avoided.

Several slip gauges are combined together temporarily to provide the end standard of a specific length. A set of slip gauges should enable the user to stack them together to provide an accuracy of up to one-thousandth of a millimetre or better. In order to achieve this, individual gauges must be available in dimensions needed to achieve any combination within the available number of gauges. The surfaces of neighbouring slip gauges should stick so close together that there should not be any scope for even a layer of air to be trapped between them, which can add error to the final reading. For this to happen, there should be absolute control over the form, flatness, parallelism, surface finish, dimensional stability of material, and homogeneity of gauging surfaces. While building slip gauges to the required height, the surfaces of slip gauges are pressed

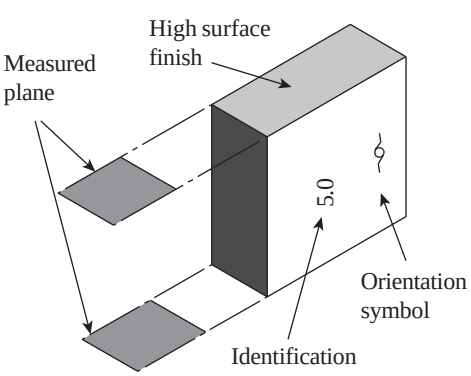


Fig. 4.42 Functional features of a slip gauge

into contact by imparting a small twisting motion while maintaining the contact pressure. The slip gauges are held together due to molecular adhesion between a liquid film and the mating surfaces. This phenomenon is known as 'wringing'.

4.9.1 Gauge Block Shapes, Grades, and Sizes

Slip gauges are available in three basic shapes: rectangular, square with a central hole, and square without a central hole. Rectangular blocks are the most commonly used since they can be used conveniently where space is restricted and excess weight is to be avoided. Square slip gauges have larger surface area and lesser wear rate because of uniform distribution of stresses during measurements. They also adhere better to each other when wrung together. Square gauge blocks with central holes permit the use of tie rods, which ensure that the built-up slip gauges do not fall apart.

Slip gauges are classified into grades depending on their *guaranteed* accuracy. The grade defines the type of application for which a slip gauge is suited, such as inspection, reference, or calibration. Accordingly, slip gauges are designated into five grades, namely grade 2, grade 1, grade 0, grade 00, and inspection grade.

Grade 2 This is the workshop-grade slip gauge. Typical uses include setting up machine tools, milling cutters, etc., on the shop floor.

Grade 1 This grade is used for tool room applications for setting up sine bars, dial indicators, calibration of vernier, micrometer instruments, and so on.

Grade 0 This is an inspection-grade slip gauge. Limited people will have access to this slip gauge and extreme care is taken to guard it against rough usage.

Grade 00 This set is kept in the standards room and is used for inspection/calibration of high precision only. It is also used to check the accuracy of the workshop and grade 1 slip gauges.

Calibration grade This is a special grade, with the actual sizes of slip gauges stated on a special chart supplied with the set of slip gauges. This chart gives the exact dimension of the slip gauge, unlike the previous grades, which are presumed to have been manufactured to a set tolerance. They are the best-grade slip gauges because even though slip gauges are manufactured using precision manufacturing methods, it is difficult to achieve 100% dimensional accuracy. Calibration-grade slip gauges are not necessarily available in a set of preferred sizes, but their sizes are explicitly specified up to the third or fourth decimal place of a millimetre.

Many other grading standards are followed for slip gauges, such as JIS B 7506-1997 (Japan), DIN 861-1980 (Germany), ASME (USA), and BS 4311:Part 1:1993 (UK). Most of these standards assign grades such as A, AA, AAA, and B. While a grade B may conform to the workshop-grade slip gauge, grades AA and AAA are calibration and reference grades, respectively.

Slip gauges are available in standard sets in both metric and inch units. In metric units, sets of 31, 48, 56, and 103 pieces are available. For instance, the set of 103 pieces consists of the following:

1. One piece of 1.005 mm
2. 49 pieces ranging from 1.01 to 1.49 mm in steps of 0.01 mm

3. 49 pieces ranging from 0.5 to 24.5 mm in steps of 0.5 mm
4. Four pieces ranging from 25 to 100 mm in steps of 25 mm

A set of 56 slip gauges consists of the following:

1. One piece of 1.0005 mm
2. Nine pieces ranging from 1.001 to 1.009 mm in steps of 0.001 mm
3. Nine pieces ranging from 1.01 to 1.09 mm in steps of 0.01 mm
4. Nine pieces ranging from 1.0 to 1.9 mm in steps of 0.1 mm
5. 25 pieces ranging from 1 to 25 mm in steps of 1.0 mm
6. Three pieces ranging from 25 to 75 mm in steps of 25 mm

Generally, the set of slip gauges will also include a pair of tungsten carbide protection gauges. These are marked with letter 'P', are 1 or 1.5 mm thick, and are wrung to the end of the slip gauge combination. They are used whenever slip gauges are used along with instruments like sine bars, which are made of metallic surfaces that may accelerate the wear of regular slip gauges. Wear blocks are also recommended when gauge block holders are used to hold a set of wrung gauges together. The purpose of using a pair of wear blocks, one at the top and the other at the bottom of the stack, is to ensure that major wear is concentrated over the two wear gauges, which can be economically replaced when worn out. This will extend the useful life of the set of slip gauges.

4.9.2 Wringing of Slip Gauges

Wringing is the phenomenon of adhesion of two flat and smooth surfaces when they are brought into close contact with each other. The force of adhesion is such that the stack of a set of blocks will almost serve as a single block, and can be handled and moved around without disturbing the position of individual blocks. More importantly, if the surfaces are clean and flat, the thin layer of film separating the blocks will also have negligible thickness. This means that stacking of multiple blocks of known dimensions will give the overall dimension with minimum error.

Wringing Phenomenon

When two surfaces are brought into contact, some amount of space exists between them. This is because of surface irregularities and presence of dirt, oil, grease, or air pockets. Let us assume that the two surfaces are perfectly flat with highly finished surfaces, free from dirt and oil, and firmly pressed together. Now the air gap becomes so small that it acts in the same way as a liquid film. The thickness of this film can be as low as 0.00001 mm. Now a question arises as to why the blocks stick together so firmly that even a high magnitude of force acting perpendicular to their surfaces will not be able to separate them. A combination of two factors appears to ensure this high adhesion force. First, as shown in Fig. 4.43, an atmospheric force of 1 bar is acting in the direction shown by the two arrows. This is contributing to the adhesion of the surfaces of the two slip gauges.

Secondly, the surfaces are in such close proximity that there is molecular adhesion of high magnitude that creates a high adhesion force. Since the slip gauge surfaces undergo lapping as a super finishing operation, material removal takes place at the molecular level. Since some molecules are lost during the lapping operation, the material is receptive to molecules of the mating surface, which creates high molecular adhesion.

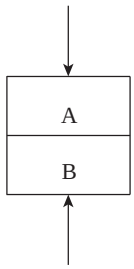


Fig. 4.43 Wringing phenomenon

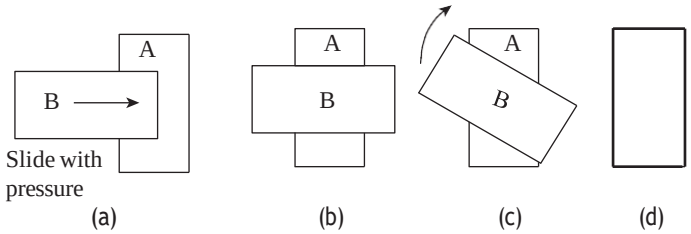


Fig. 4.44 Technique of wringing slip gauges (a) Step 1 (b) Step 2 (c) Step 3 (d) Step 4

These two factors collectively ensure adhesion of slip gauges with minimum air gap between them. Therefore, a stack of slip gauges will have a length equal to the sum of the individual heights.

Technique of Wringing Slip Gauges

The ability of a given gauge block to wring is called *wringability*; it is defined as 'the ability of two surfaces to adhere tightly to each other in the absence of external means'. The minimum conditions for wringability are a surface finish of 0.025 µm or better, and a flatness of at least 0.13 µm. Wringing of slip gauges should be done carefully and methodically because a film of dust, moisture, or oil trapped between gauges will reduce the accuracy of measurement. The first step is to clean the slip gauges immediately before wringing because any gap in time will allow dust and moisture to settle on the gauges. A very fine hairbrush can be used to clean them. Some people are under the false notion that a thin film of oil should always be applied to the gauge surfaces before wringing. Most often, the application of oil itself may introduce unwanted dust and oil in between the gauges. The need for additional oil film is felt for worn out gauges where there is reduced metal-to-metal contact resulting in poor molecular adhesion. The following are the preferred steps in the wringing of slip gauges:

1. Clean slip gauge surfaces with a fine hairbrush (camel hairbrushes are often recommended) and a dry pad.
2. Overlap gauging surfaces by about one-fourth of their length, as shown in Fig. 4.44(a).
3. Slide one block perpendicularly across the other by applying moderate pressure. The two blocks should now form the shape as shown in Fig. 4.44(b).
4. Now, gently rotate one of the blocks until it is in line with the other block, as in Fig. 4.44(c) and (d).

Combining Slip Gauges

As pointed out earlier, gauge blocks are available in standard sets of 31, 48, 56, and 103 pieces. While figuring out the slip gauges that are required to make up a given dimension, a procedure must be followed to save time and, more importantly, to ensure that a minimum number of gauges are used. Please remember that more the number of gauges used, more is the separation of gauges by a thin film, which can cumulatively contribute to substantial error. In addition, the accuracy up to which a dimension can be built depends on the gauge that can give accuracy up to the last possible decimal place. For instance, while the 103-piece set can give an accuracy of up to 0.005 mm, the 56-piece set can give up to 0.005 mm.

ANGULAR MEASUREMENTS

After studying this chapter, the reader will be able to

- understand the basic requirements of angular measurement in the industry and the variety of instruments available at our disposal
- elucidate the basic principle of a protractor and its extension as the universal bevel protractor, which is an indispensable part of a metrology laboratory
- measure angles using the sine principle and explain the use of sine bar, sine block, sine plate, and sine centre
- use angle gauges to set them accurately to the required angle
- appreciate the importance of 'bubble instruments' such as the conventional spirit level and clinometers in angular measurement
- explain the principles of optical measurement instruments, the most popular ones being the autocollimator and the angle dekkor

5.1 INTRODUCTION

Length standards such as foot and metre are arbitrary inventions of man. This has necessitated the use of wavelength of light as a reference standard of length because of the difficulty in accurately replicating the earlier standards. On the other hand, the standard for angle, which is derived with relation to a circle, is not man-made but exists in nature. One may call it degree or radian, but the fact remains that it has a direct relation to a circle, which is an envelope of a line moving about one of its ends. Whether one defines a circle as the circumference of a planet or path of an electron around the nucleus of an atom, its parts always bear a unique relationship.

The precise measurement of angles is an important requirement in workshops and tool rooms. We need to measure angles of interchangeable parts, gears, jigs, fixtures, etc. Some of the typical measurements are tapers of bores, flank angle and included angle of a gear, angle made by a seating surface of a jig with respect to a reference surface, and taper angle of a jib. Sometimes, the primary objective of angle measurement is not to measure angles. This may sound rather strange, but this is the case in the assessment of alignment of machine parts. Measurement of straightness, parallelism, and flatness of machine parts requires highly

sensitive instruments like autocollimators. The angle reading from such an instrument is a measure of the error of alignment.

There are a wide range of instruments, starting from simple scaled instruments to sophisticated types that use laser interferometry techniques. The basic types are simple improvisations of a protractor, but with better discrimination (least count), for example, a vernier protractor. These instruments are provided with a mechanical support or a simple mechanism to position them accurately against the given workpiece and lock the reading. A spirit level has universal applications, not only in mechanical engineering but also in civil engineering construction for aligning structural members such as beams and columns. Instruments employing the basic principle of a spirit level but with higher resolution, such as conventional or electronic clinometers, are popular in metrology applications. By far, the most precise instruments are collimators and angle dekkors, which belong to the family of instruments referred to as *optical tooling*. This chapter deals with some of the popular angle measurement devices that are widely used in the industry.

5.2 PROTRACTOR

A simple protractor is a basic device used for measuring angles. At best, it can provide a least count of 1° for smaller protractors and ½° for large ones. However, simple though it may be, the user should follow the basic principles of its usage to measure angles accurately. For instance, the surface of the instrument should be parallel to the surface of the object, and the reference line of the protractor should coincide perfectly with the reference line of the angle being measured. Positioning of the protractor and observation of readings should be performed with care to avoid parallax error.

Similar to a steel rule, a simple protractor has limited usage in engineering metrology. However, a few additions and a simple mechanism, which can hold a main scale, a vernier scale, and a rotatable blade, can make it very versatile. A universal bevel protractor is one such instrument that has a mechanism that enables easy measurement and retention of a reading. A vernier scale improves the least count substantially. Additional attachments enable the measurement of acute and obtuse angles with ease and thereby justify its name as the universal bevel protractor. It can measure the angle enclosed by bevelled surfaces with ease and hence the name.

In fact, if one traces the history of development of angle-measuring devices, the bevel protractor preceded the universal bevel protractor. The earliest bevel protractor had a simple mechanism that facilitated rotation of measuring blades and locked them in place. It had a scale graduated in degrees on which the measurements could be directly read. However, these instruments have largely been replaced by universal bevel protractors and the older types are not being used in metrology applications now. Therefore, we shall directly go to a discussion on the universal bevel protractor.

5.2.1 Universal Bevel Protractor

The universal bevel protractor with a 5' accuracy is commonly found in all tool rooms and metrology laboratories. Figure 5.1 illustrates the construction of a universal bevel protractor. It has a base plate or stock whose surface has a high degree of flatness and surface finish. The stock is placed on the workpiece whose angle is to be measured. An adjustable blade attached

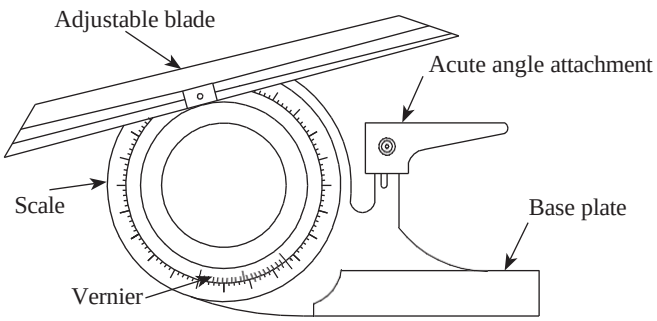


Fig. 5.1 Universal bevel protractor

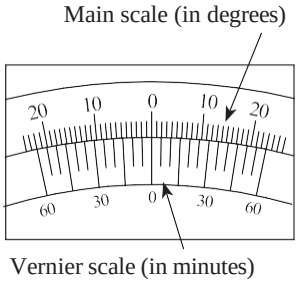


Fig. 5.2 Divisions on the vernier scale

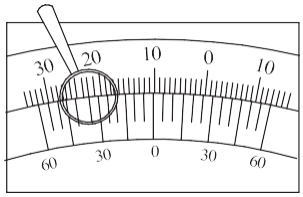


Fig. 5.3 Reading the vernier scale

to a circular dial is made to coincide with the angular surface. It can be swivelled to the required angle and locked into position to facilitate accurate reading of the circular scale that is mounted on the dial. The main scale on the dial is graduated in degrees and rotates with the rotation of the adjustable blade. A stationary vernier scale mounted close to the dial, as shown in Fig. 5.1, enables measurements to a least count of 5' or less. An acute angle attachment is provided for the measurement of acute angles.

The main scale on the dial is divided into four quadrants, each measuring 90°. Each division on this scale reads 1°. The degrees are numbered from 0 to 90 on either side of the zeroth division. The vernier scale has 24 divisions, which correspond to 46 divisions on the main scale. However, the divisions on the vernier scale are numbered from 0 to 60 on either side of the zeroth division, as shown in Fig. 5.2.

Calculation of Least Count

Value of one main scale division = 1°

24 vernier divisions correspond to 46 main scale divisions. From Fig. 5.2, it is clear that one vernier division equals 1/12th of 23°. Let us assume that the zeroth division on both the main and the vernier scales are lined up to coincide with each other. Now, as the dial rotates, a vernier division, starting from the fifth minute up to the 60th minute, progressively coincides with a main scale division until the zeroth division on the vernier scale moves over the main scale by 2°.

Therefore, the least count is the difference between one vernier division and two main scale divisions, which is 1/12° or 5'.

Reading Vernier Scales

Consider the situation shown in Fig. 5.3. The zeroth division of the vernier scale is just past the 10° division on the main scale. The seventh division, marked as the 35' division, on the left-hand side of the vernier scale coincides with a division on the main scale. Therefore, the reading in this case is 10°35'.

Sometimes, confusion arises regarding the direction in which the vernier has to be read. This confusion may crop up for the aforementioned example also. It is possible that a division on the right side of zero on the vernier scale may be coinciding with a division on the main scale (dial scale). In order to eliminate this confusion, we follow a simple rule. *Always read the*

vernier from zero in the same direction that you read the dial scale. In the given example, the 10th division on the dial, which is close to the zeroth division on the vernier, is to the left of the zeroth division on the dial scale. In other words, the dial scale is being read in the leftward or anticlockwise direction. Therefore, the vernier should also be read towards the left of the vernier zero division. Figure 5.4 illustrates the use of a bevel protractor for measurement of angles. While Fig. 5.4(a) illustrates the use of acute angle attachment, Fig. 5.4(b) shows how the angle of an inside bevelled face can be measured.

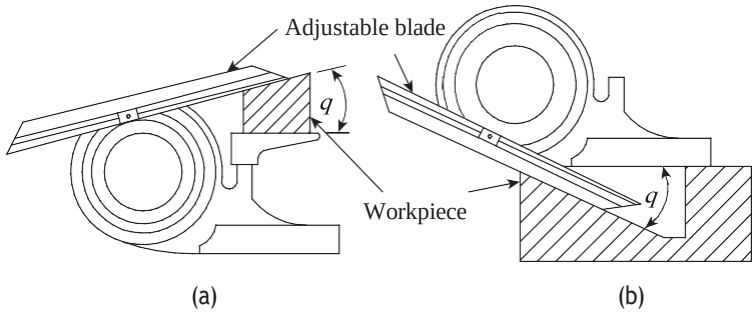


Fig. 5.4 Measurement of angles using bevel protractor (a) Acute angle attachment (b) Inside bevelled face angle measurement

Angles and their Supplements

Since a universal bevel protractor can measure both acute and obtuse angles, care should be exercised to clearly differentiate between the angle being indicated on the scale and its supplement. The dial gauge is graduated from 0° to 90° in four quadrants, as shown in Fig. 5.5. Figure 5.5(a) illustrates the orientation of the blade with respect to the base when the protractor is set to 90°. The zeroth division on the vernier coincides with the 90° division on the dial scale.

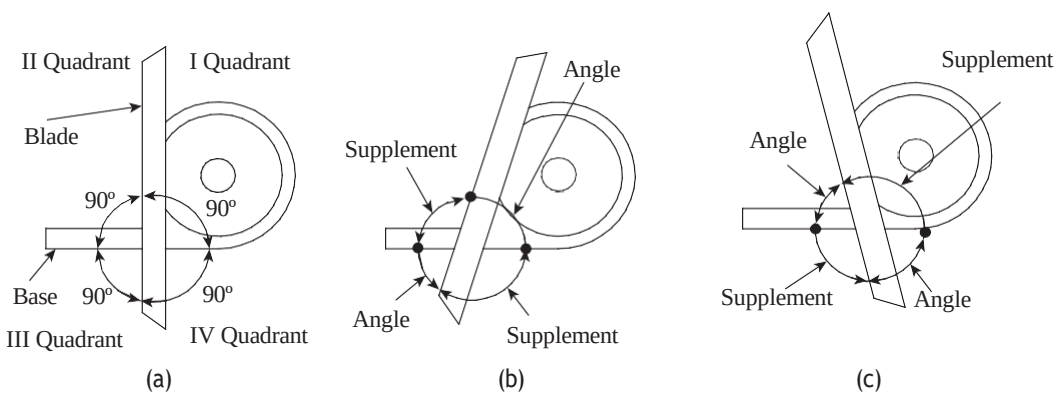


Fig. 5.5 Angles and their supplements (a) Blade oriented with base (b) Blade turned clockwise (c) Blade turned counterclockwise

Suppose the blade is turned clockwise, as in Fig. 5.5(b), the angles read directly are those that are formed from the blade to the base counterclockwise. Thus, if the angles of a work part are being measured in quadrant I or III, the angles can be read directly from the scale. On the other hand, if the angles of a work part are being measured in quadrant II or IV, the actual angles are given by their supplements. In other words, the value of the angle is obtained by subtracting the angle indicated by the scale from 180°. Both these angles, obviously, are obtuse angles.

Further, Fig. 5.5(c) illustrates a situation when the blade is turned counterclockwise. Here also, angles can be directly read only in two quadrants, namely the second and the fourth. These angles are formed in the clockwise direction from the blade to the base, and are acute angles. The supplements in I and III quadrants give the obtuse angles of the work parts, which are held in these quadrants.

Bevel protractors in general are classified into four types: A, B, C, and D. Types C and D are the basic types, with the dial scale graduated in degrees. They are neither provided with a vernier scale nor a fine adjustment device. Types A and B are provided with a vernier scale, which can accurately read up to 5'. While type A has both an acute angle attachment and a fine adjustment device, type B does not have either of them.

A bevel protractor is a precision angle-measuring instrument. To ensure an accurate measurement, one should follow these guidelines:

1. The instrument should be thoroughly cleaned before use. It is not recommended to use compressed air for cleaning, as it can drive particles into the instrument.
2. It is important to understand that the universal bevel protractor does not essentially measure the angle on the work part. It measures the angle between its own parts, that is, the angle between the base plate and the adjustable blade. Therefore, one should ensure proper and intimate contact between the protractor and the features of the part.
3. An easy method to determine if the blade is in contact with the work part is to place a light behind it and adjust the blade so that no light leaks between the two.
4. It should always be ensured that the instrument is in a plane parallel to the plane of the angle. In the absence of this condition, the angle measured will be erroneous.
5. The accuracy of measurement also depends on the surface quality of the work part. Burrs and excessive surface roughness interfere with the intimate contact between the bevel protractor and the work part, leading to erroneous measurements.
6. One should be careful to not slide the instrument over hard or abrasive surfaces, and not over-tighten clamps.
7. Before replacing the instrument in its case, it has to be wiped with a clean and dry cloth, a thin rust-preventing coating has to be applied, and moving parts need to be lubricated.

5.2.2 Optical Bevel Protractor

An optical protractor is a simple extension of the universal bevel protractor. A lens in the form of an eyepiece is provided to facilitate easy reading of the protractor scale. Figure 5.6 illustrates the construction details of an optical bevel protractor. The blade is clamped to the dial by means of a blade clamp. This enables fitting of blades of different lengths, depending on the work part being measured. In a protractor without a vernier, the dial scale reading can be

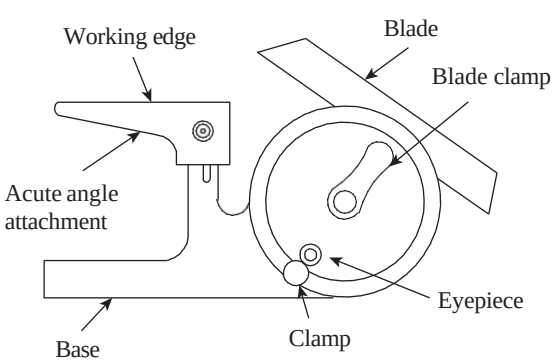


Fig. 5.6 Optical bevel protractor

directly read through the eyepiece. In vernier protractors, the eyepiece is attached on top of the vernier scale itself, which together move as a single unit over the stationary dial scale. The eyepiece provides a magnified view of the reading for the convenience of the user.

Most of the universal protractors in use are of this type. An acute angle attachment is provided to facilitate measurement of acute angles on work parts. A clamp is provided to lock the reading, so that it can be read and recorded at a convenient position by the user.

5.3 SINE BAR

A sine bar is used to measure angles based on the sine principle. Its upper surface forms the hypotenuse of a triangle formed by a steel bar terminating in a cylinder near each end. When one of the cylinders, called a roller, is resting on a flat surface, the bar can be set at any desired angle by simply raising the second cylinder. The required angle is obtained when the difference in height between the two rollers is equal to the sine of the angle multiplied by the distance between the centres of the rollers. Figure 5.7 illustrates the construction details of a sine bar.

Sine bars are made of corrosion-resistant steel, and are hardened, ground, and stabilized. The size is specified by the distance between the centres of the cylinders, which is 100, 200, or 300 mm. The upper surface has a high degree of flatness of up to 0.001 mm for a 100 mm length and is perfectly parallel to the axis joining the centres of the two cylinders. The parallelism of upper surface with the datum line is of the order of 0.001 mm for a 100 mm length. Relief holes are sometimes provided to reduce the weight of the sine bar. This by itself is not a complete measuring instrument. Accessories such as a surface plate and slip gauges are needed to perform the measurement process. Figure 5.8 illustrates the application of a sine rule for angle measurement.

The sine of angle q formed between the upper surface of a sine bar and the surface plate (datum) is given by

$$\text{Sin}\,(q) = h/L$$

where h is the height difference between the two rollers and L is the distance between the centres of the rollers.

Therefore, $h = L\ \text{Sin}\,(q)$

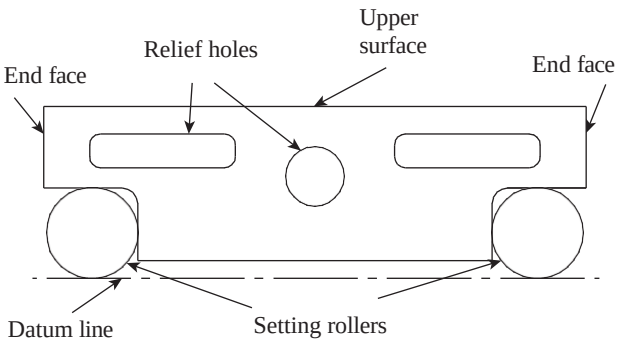


Fig. 5.7 Sine bar

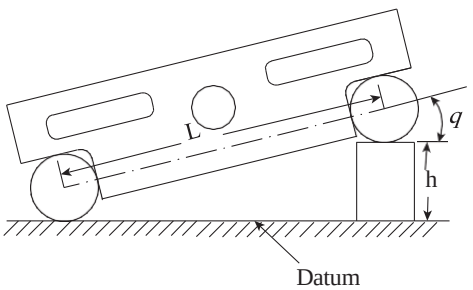


Fig. 5.8 Sine rule

5.3.1 Setting Sine Bars to Desired Angles

By building slip gauges to height *h* and placing the sine bar on a surface plate with one roller on top of the slip gauges, the upper surface can be set to a desired angle with respect to the surface plate. The set-up is easier because the cylinders are integral parts of the sine bar and no separate clamping is required. No measurement is required between the cylinders since this is a known length. It is preferable to use the sine bar on a grade A surface plate. In addition, it is desirable to support both rollers on gauge blocks so that the minute irregularities of the surface plate may be eliminated.

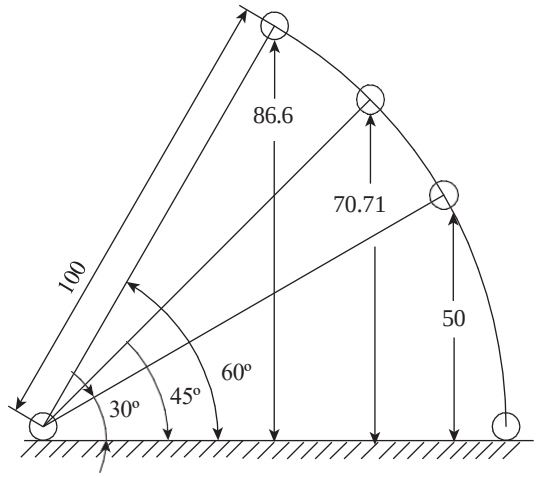


Fig. 5.9 Change in sine value with change in angle

The question often asked is about the maximum angle that can be set using a sine bar. The standard response is 45°. At higher angles, errors due to the distance between the centres of the rollers and gauge blocks get magnified. This is explained in the following example.

Let us assume a 100 mm-long sine bar and calculate the heights to set it at 30°, 45°, and 60°, as shown in Fig. 5.9. The heights are respectively 53, 73.71, and 89.6 mm for angles 30°, 60°, and 90°.

Assume there is an error of +0.1 mm in the height, *h*. Table 5.1 illustrates the error in the measured angles. The actual angle in 'degrees' is given by $\sin^{-1}(h/L)$.

Table 5.1 Relationship between error and angle being set

Angle to be set (degrees)	Length of sine bar (mm)	Height of slip gauges (mm)	Actual angle (degrees)	Error in measurement (degrees)
30	100.0	50	30	0.066
		50.1	30.066	
45	100.0	70.71	45	0.080
		70.81	45.080	
60	100.0	86.6	60	0.112
		86.7	60.112	

Accounting for an error of 0.1 mm in height *h*, the angular error for 30° is 0.066 mm. This error increases to 0.08 mm for 45° and jumps to 0.112 mm for 60°. A similar increase in error is observed if the distance between the centres of the rollers of the sine bar has an error. This is the primary reason why metrologists avoid using sine bars for angles greater than 45°.

5.3.2 Measuring Unknown Angles with Sine Bar

A sine bar can also be used to measure unknown angles with a high degree of precision. The angle of the work part is first measured using an instrument like a bevel protractor. Then, the work part is clamped to the sine bar and set on top of a surface plate to that angle using slip gauges, as shown in Fig. 5.10 (clamping details are not shown in the figure).

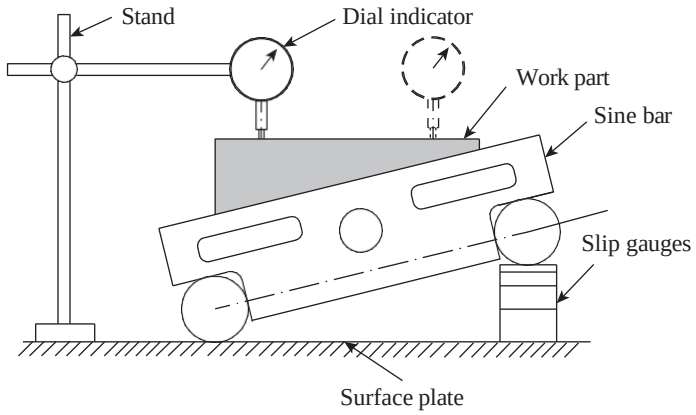


Fig. 5.10 Measurement of unknown angles

A dial gauge fixed to a stand is brought in contact with the top surface of the work part at one end and set to zero. Now, the dial indicator is moved to the other end of the work part in a straight line. A zero reading on the dial indicator indicates that the work part surface is perfectly horizontal and the set angle is the right one. On the other hand, if the dial indicator shows any deviations, adjustments in the height of slip gauges is necessary to ensure that the work part surface is horizontal. The difference in height corresponding to the dial gauge reading is incorporated in the slip gauges, and the procedure is repeated until the dial indicators show zero deviation. The actual angle is calculated using the total height of the slip gauges.

Instead of a dial gauge, a high-amplification comparator can be used for better accuracy. Whether setting a sine bar to a known angle or for measuring unknown angles, a few guidelines should be followed to ensure proper usage of the instrument:

1. It is not recommended to use sine bars for angles greater than 45° because any error in the sine bar or height of slip gauges gets accentuated.
2. Sine bars provide the most reliable measurements for angles less than 15°.
3. The longer the sine bar, the better the measurement accuracy.
4. It is preferable to use the sine bar at a temperature recommended by the supplier. The accuracy of measurement is influenced by the ambient temperature.
5. It is recommended to clamp the sine bar and the work part against an angle plate. This prevents misalignment of the workpiece with the sine bar while making measurements.
6. One should always keep in mind that the sine principle can be put to use provided the sine bar is used along with a high-quality surface plate and set of slip gauges.

5.3.3 Sine Blocks, Sine Plates, and Sine Tables

A *sine block* is a sine bar that is wide enough to stand unsupported (Fig. 5.11). If it rests on an integral base, it becomes a *sine plate* (Fig. 5.12). A sine plate is wider than a sine block. A heavy-duty sine plate is rugged enough to hold work parts for machining or inspection of angles. If a sine plate is an integral part of another device, for example, a machine tool, it is called a *sine table*. However, there is no exact dividing line between the three.



Fig. 5.11 Sine block

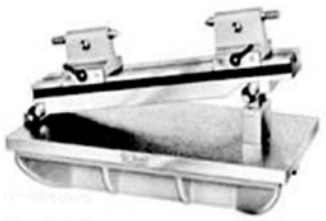


Fig. 5.12 Sine plate

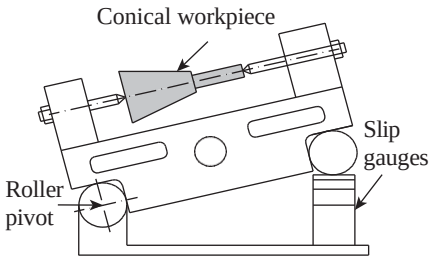


Fig. 5.13 Sine centre

In all these three devices, the work part rests on them. They are often used like a fixture to keep the work part at a particular orientation, so that the required angle is machined. The instruments have attachments to raise and lock the block to the required angle, and to also fasten work parts. The sine table is the most rugged device, which may be swung to any angle from 0° to 90° by pivoting about the hinged end.

There are many instances where compound angles need to be machined or inspected. While simple angles lie on one plane, compound angles of a surface lie on more than one plane. In a surface formed by the intersections of planes, the angles on the surface planes are called *face angles*. A compound sine plate can conveniently measure or set itself to this face angle. In a typical compound sine plate, there are two sine plates: a base plate creates one plane, while the top plate creates the second plane. Compound sine plates are usually used for finishing operations, for example, a finish grinding operation.

5.3.4 Sine Centre

A sine centre provides a convenient means of measuring angles of conical workpieces that are held between centres, as shown in Fig. 5.13. One of the rollers is pivoted about its axis, thereby allowing the sine bar to be set to an angle by lifting the other roller. The base of the sine centre has a high degree of flatness, and slip gauges are wrung and placed on it to set the sine bar at the required angle.

Conical workpieces that need to be inspected are placed between the centres. The sine centre is used for measuring angles up to 60°. The procedure for measuring angles is very similar to the one described in Section 5.3.2. A dial gauge clamped to a stand is set against the conical workpiece. The sine bar is set to an angle such that the dial gauge registers no deviation when moved from one end of the workpiece to the other. The angle is determined by applying the sine rule.

5.4 ANGLE GAUGES

Angle gauges, which are made of high-grade wear-resistant steel, work on a principle similar to slip gauges. While slip gauges can be built to give linear dimensions, angle gauges can be built to give the required angle. The gauges come in a standard set of angle blocks that can be wrung together in a suitable combination to build an angle. C.E. Johansson, who developed slip gauges, is also credited with the invention of angle gauge blocks. However, the first set of a combination of angle gauges was devised by Dr G.A. Tomlinson of the National Physical

Laboratory, UK, in the year 1939, which provided the highest number of angle combinations. His set of 10 blocks can be used to set any angle between 0° and 180° in increments of 5'.

At the outset, it seems improbable that a set of 10 gauges is sufficient to build so many angles. However, angle blocks have a special feature that is impossible in slip gauges—the former can be subtracted as well as added. This fact is illustrated in Fig. 5.14.

This illustration shows the way in which two gauge blocks can be used in combination to generate two different angles. If a 5° angle block is used along with a 30° angle block, as shown in Fig. 5.14(a), the resulting angle is 35°. If the 5° angle block is reversed and combined with the 30° angle block, as shown in Fig. 5.14(b), the resulting angle is 25°. Reversal of an angle block subtracts itself from the total angle generated by combining other angle blocks. This provides the scope for creating various combinations of angle gauges in order to generate angles that are spread over a wide range by using a minimum number of gauges.

Angle gauges are made of hardened steel, which is lapped and polished to a high degree of accuracy and flatness. The gauges are about 75 mm long and 15 mm wide, and the two surfaces that generate the angles are accurate up to ±2". The gauges are available in sets of 6, 11, or 16. Table 5.2 provides the details of individual blocks in these sets.

Most angles can be combined in several ways. However, in order to minimize error, which gets compounded if the number of gauges used is increased, it is preferable to use the least number of angle gauge blocks. The set of 16 gauges forms all the angles between 0° and 99° in 1" steps—a total of 3,56,400 combinations! The laboratory master-grade set has an accuracy of one-fourth of a second. While the inspection-grade set has an accuracy of ½", the tool room-grade set has an accuracy of 1".

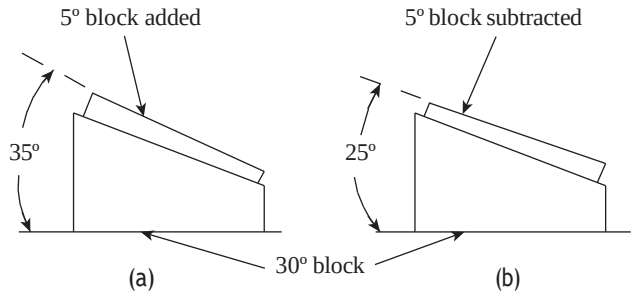


Fig. 5.14 Angle gauge block (a) Addition (b) Subtraction

Table 5.2 Angle gauge block sets

Smallest increment by which any angle can be produced	Number of individual blocks contained in the set	Detailed listing of the blocks composing the set
1°	6	Six blocks of 1°, 3°, 5°, 15°, 30°, and 45°
1′	11	Six blocks of 1°, 3°, 5°, 15°, 30°, and 45° Five blocks of 1', 3', 5', 20', and 30′
1″	16	Six blocks of 1°, 3°, 5°, 15°, 30°, and 45° Five blocks of 1', 3', 5', 20', and 30' Five blocks of 1", 3", 5", 20", and 30"

PART II

Mechanical Measurements

- **Measurement Systems**
- **Transducers**
- **Measurement of Force, Torque, and Strain**
- **Measurement of Temperature**
- **Pressure Measurements**

CHAPTER-3

Measurement Systems

After studying this chapter, the reader will be able to

- understand the effect of hysteresis in measuring instruments
- appreciate the importance of linearity
- explain some basic definitions in measurements
- appreciate the significance of system response
- elucidate the different functional elements of generalized measurement systems

12.1 INTRODUCTION

We know that measurement is defined as the quantification of a physical variable using a measuring instrument. During the process of measurement, a specific value is assigned to the unknown quantity after due comparison with a predefined standard. The measuring process is schematically represented in Fig. 12.1.

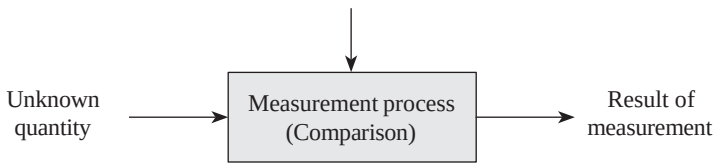


Fig. 12.1 Measuring process

t is pertinent to mention here that while making physical measurements, one needs to keep in mind that measurements are not completely accurate, as discussed in Chapter 1. It has to be remembered that the act of measurement will be complete when the uncertainty associated with measurements is also considered. In order to understand the uncertainty associated with measurement, we need to study a few more definitions, apart from those discussed in the first chapter. It is important to know various measurement attributes that affect the performance of a measuring instrument. The operating parameters within which an instrument performs the measurement process needs to be understood clearly.

12.2 SOME BASIC DEFINITIONS

In order to clearly understand the concepts of measurement, we will now look into a few more definitions.

12.2.1 Hysteresis in Measurement Systems

When the value of the measured quantity remains the same irrespective of whether the measurements have been obtained in an ascending or a descending order, a system is said to be free from hysteresis. Many instruments do not reproduce the same reading due to the presence of hysteresis. Slack motion in bearings and gears, storage of strain energy in the system, bearing friction, residual charge in electrical components, etc., are some of the reasons for the occurrence of hysteresis. Figure 12.2 shows a typical hysteresis loop for a pressure gauge. If the width of the hysteresis band formed is appreciably more, the average of the two measurements (obtained in both ascending and descending orders) is used. However, the presence of some hysteresis in measuring systems is normal, and the repeatability of the system is affected by it.

12.2.2 Linearity in Measurement Systems

It is desirable to design instruments having a linear relationship between the applied static input and the indicated output values, as shown in Fig. 12.3. A measuring instrument/system is said to be linear if it uniformly responds to incremental changes, that is, the output value is equal to the input value of the measured property over a specified range.

Linearity is defined as the maximum deviation of the output of the measuring system from a specified straight line applied to a plot of data points on a curve of measured (output) values versus the measurand (input) values. In order to obtain accurate measurement readings, a high degree of linearity should be maintained in the instrument or efforts have to be made to minimize linearity errors. A better degree of linearity renders the instrument to be readily calibrated. However, in practice, only an approximation of the linearity is achieved as there is always some small variance associated with the measuring system. Hence, the expected linearity of the input is usually specified as a percentage of the operating range.

A fuel gauge in an automobile is one such example of non-linearity. When the tank is completely filled with fuel, the gauge needle indicates a full tank. The needle continues to be in the nearly full position even after the consumption of a significant amount of fuel. However, after considerable travel, the needle appears to move rapidly towards the minimum fuel value.

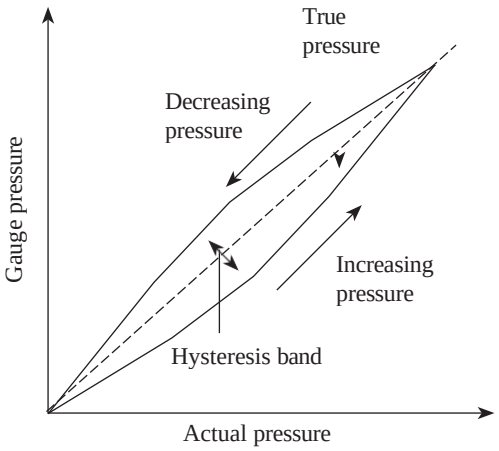


Fig. 12.2 Hysteresis in measurement systems

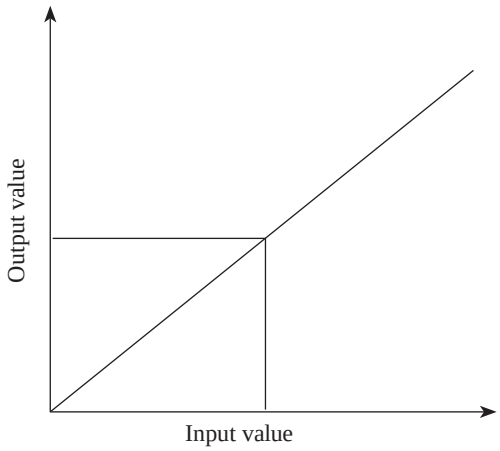


Fig. 12.3 Linearity in measurements

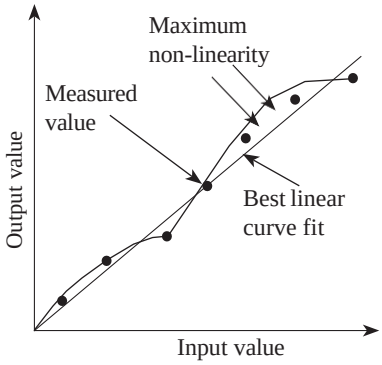


Fig. 12.4 Linearity errors

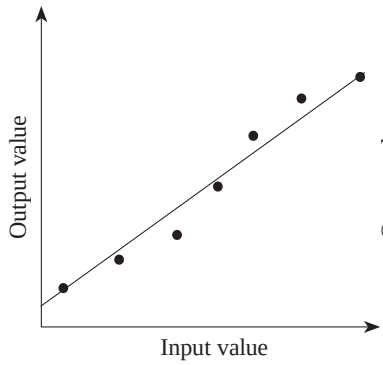


Fig. 12.5 End point line

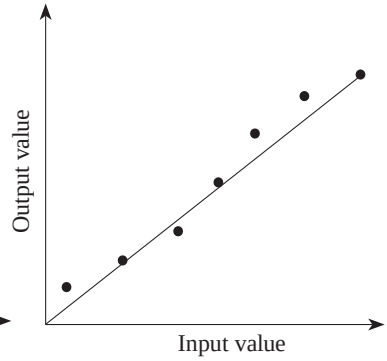


Fig. 12.6 Terminal line

This may be due to the improper calibration of the fuel gauge at the maximum and minimum values of the operating range.

Before making any interpretation or comparison of the linearity specifications of the measuring instrument, it is necessary to define the exact nature of the reference straight line adopted, as several lines can be used as the reference of linearity. The most common lines are as follows:

Best-fit line The plot of the output values versus the input values with the best line fit is shown in Fig. 12.4. The line of best fit is the most common way to show the correlation between two variables. This line, which is also known as the trend line, is drawn through the centre of a group of data points on a scatter plot. The best-fit line may pass through all the points, some of the points, or none of the points.

End point line This is employed when the output is bipolar. It is the line drawn by joining the end points of the data plot without any consideration of the origin. This is represented in Fig. 12.5.

Terminal line When the line is drawn from the origin to the data point at full scale output, it is known as terminal line. The terminal line is shown in Fig. 12.6.

Least square line This is the most preferred and extensively used method in regression analysis. Carl Freidrich Gauss (1975) was the first to give the earliest description of the least square method. It is a statistical technique and a more precise way of determining the line of best fit for a given set of data points. The best-fit line is drawn through a number of data points by minimizing the sum of the squares of the deviations of the data points from the line of best fit, hence the name least squares. The line is specified by an equation relating the input value to the output value by considering the set of data points.

12.2.3 Resolution of Measuring Instruments

Resolution is the smallest change in a physical property that an instrument can sense. For example, a weighing machine in a gymnasium normally senses weight variations in kilograms, whereas a weighing machine in a jewellery shop can detect weight in milligrams. Naturally, the weighing machine in the jewellery shop has a superior resolution than the one at the gymnasium.

Resolution of an instrument can also be defined as the minimum incremental value of the input signal that is required to cause a detectable change in the output.

12.2.4 Threshold

If the input to the instrument is gradually increased from zero, a minimum value of that input is required to detect the output. This minimum value of the input is defined as the threshold of the instrument. The numerical value of the input to cause a change in the output is called the threshold value of the instrument.

12.2.5 Drift

Drift can be defined as the variation caused in the output of an instrument, which is not caused by any change in the input. Drift in a measuring instrument is mainly caused by internal temperature variations and lack of component stability. A change in the zero output of a measuring instrument caused by a change in the ambient temperature is known as thermal zero shift. Thermal sensitivity is defined as the change in the sensitivity of a measuring instrument because of temperature variations. These errors can be minimized by maintaining a constant ambient temperature during the course of a measurement and/or by frequently calibrating the measuring instrument as the ambient temperature changes.

12.2.6 Zero Stability

It is defined as the ability of an instrument to return to the zero reading after the input signal or measurand comes back to the zero value and other variations due to temperature, pressure, vibrations, magnetic effect, etc., have been eliminated.

12.2.7 Loading Effects

Any measuring instrument generally consists of different elements that are used for sensing, conditioning, or transmitting purposes. Ideally, when such elements are introduced into the measuring system, there should not be any distortion in the original signal. However, in practice, whenever any such element is introduced into the system, some amount of distortion occurs in the original signal, making an ideal measurement impossible. The distortion may result in wave form distortion, phase shift, and attenuation of the signal (reduction in magnitude); sometimes, all these undesirable features may combine to affect the output of the measurement.

Hence, loading effect is defined as the incapability of a measuring system to faithfully measure, record, or control the measurand in an undistorted form. It may occur in any of the three stages of measurement or sometimes it may be carried right down to the basic elements themselves.

12.2.8 System Response

One of the essential characteristics of a measuring instrument is to transmit and present faithfully all the relevant information included in the input signal and exclude the rest. We know that during measurements, the input rapidly changes with time, and hence, the output. The behaviour of the measuring system under the varying conditions of input with respect to time is known as the dynamic response.

There are two types of dynamic inputs: steady-state periodic quantity and transient magnitude. The magnitude of the steady-state periodic quantity has a definite repeating time cycle, whereas the time variation of the transient magnitude does not repeat. In some measurement applications, for the system to attain a steady state, enough time is available. In such situations, the transient characteristics of the system are not a matter of concern. In certain measurement systems, it is necessary to examine the transient behaviour of the physical variable under consideration. The design of the measurement system becomes more complicated when transient characteristics are considered.

When an input is given to a measuring instrument, a certain amount of time elapses before it indicates an output. This is because the measurement systems comprise one or more storage elements such as electrical inductance and capacitance, thermal and fluid capacitance, mass, and inertia. When an input is given to the measuring instrument, the system will not respond immediately because the energy storage elements do not allow a sudden flow of energy. The measuring instrument goes through a transient state before it finally attains a steady-state position. The following are the dynamic characteristics of a measurement system:

Speed of response One of most important characteristic of the measuring instrument, speed of response is defined as the speed with which the measuring instrument responds to the changes in the measured quantity. Some delay or lag is always associated with the measuring instrument, as it does not respond to the input instantaneously.

Measuring lag It is the time when an instrument begins to respond to a change in the measured quantity. This lag is normally due to the natural inertia of the measuring system. Measuring lag is of two types:

Retardation type In this case, the measurement system instantaneously begins to respond after the changes in the input have occurred.

Time delay type In this type, the measuring system begins to respond after a dead time to the applied input. Dead time is defined as the time required by the measuring system to begin its response to a change in the quantity to be measured. Dead time simply transfers the response of the system along the time scale, thereby causing a dynamic error. This type of measurement lag can be ignored as they are very small and are of the order of a fraction of a second. If the variation in the measured quantity occurs at a faster rate, the dead time will have an adverse effect on the performance of the system.

Fidelity It is defined as the degree to which a measurement system indicates the changes in the measured quantity without any dynamic error.

Dynamic error It is also known as a measurement error. It can be defined as the difference between the true value of a physical quantity under consideration that changes with time and the value indicated by the measuring system if no static error is assumed.

It is to be noted here that speed of response and fidelity are desirable characteristics, whereas measurement lag and dynamic error are undesirable.

12.3 FUNCTIONAL ELEMENTS OF MEASUREMENT SYSTEMS

We know that some physical quantities such as length and mass can be directly measured using measuring instruments. However, the direct measurement of physical quantities such

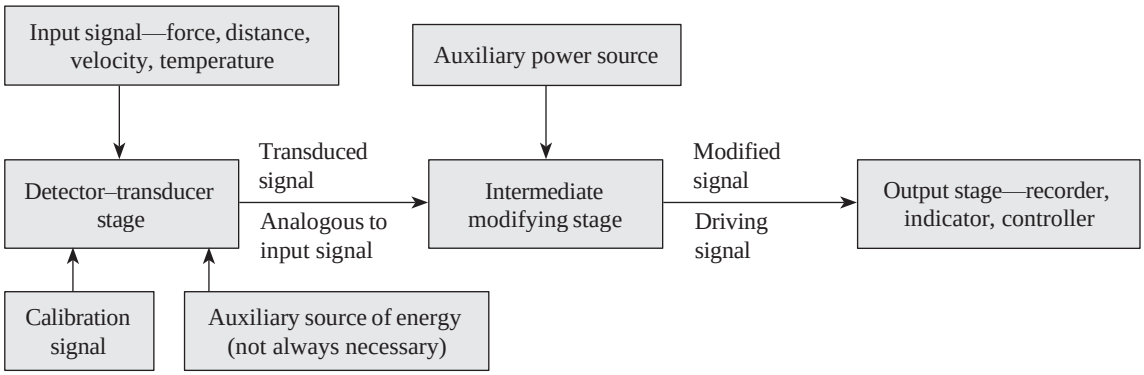


Fig. 12.7 Elements of a generalized measurement system

as temperature, force, and pressure is not possible. In such situations, measurements can be performed using a transducer, wherein one form of energy/signal that is not directly measurable is transformed into another easily measurable form. Calibration of the input and output values needs to be carried out to determine the output for all vales of input.

A measuring instrument essentially comprises three basic physical elements. Each of these elements is recognized by a functional element. Each physical element in a measuring instrument consists of a component or a group of components that perform certain functions in the measurement process. Hence, the measurement system is described in a more generalized method. A generalized measurement system essentially consists of three stages. Each of these stages performs certain steps so that the value of the physical variable to be measured is displayed as an output for our reference. Figure 12.7 schematically represents the generalized measurement systems. The three stages of a measurement system are as follows:

1. Primary detector–transducer stage
2. Intermediate modifying stage
3. Output or terminating stage

The primary detector–transducer stage senses the quantity to be measured and converts it into analogous signals. It is necessary to condition or modify the signals obtained from the primary detector–transducer stage so that it is suitable for instrumentation purposes. This signal is passed on to the intermediate modifying stage, wherein they are amplified so that they can be used in the terminating stage for display purposes. These three stages of a measurement system act as a bridge between the input given to the measuring system and its output.

12.4 PRIMARY DETECTOR–TRANSDUCER STAGE

The main function of the primary detector–transducer stage is to sense the input signal and transform it into its analogous signal, which can be easily measured. The input signal is a physical quantity such as pressure, temperature, velocity, heat, or intensity of light. The device used for detecting the input signal is known as a transducer or sensor. The transducer converts the sensed input signal into a detectable signal, which may be electrical, mechanical, optical, thermal, etc. The generated signal is further modified in the second stage. The transducer should have the ability to detect only the input quantity to be measured and exclude all other signals. The sensing process is schematically represented in Fig. 12.8.

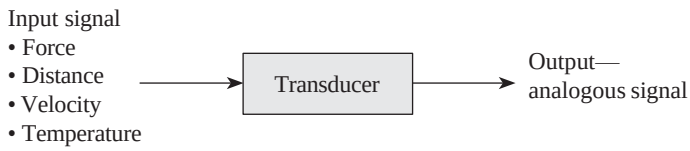


Fig. 12.8 Sensing process

For example, if bellows are employed as transducers for pressure measurement, it should detect signals pertaining to only pressure and other irrelevant input signals or disturbances should not be sensed. However, in practice, the transducers used are seldom sensitive to only the signals of the quantity being measured.

12.5 INTERMEDIATE MODIFYING STAGE

In the intermediate modifying stage of a measurement system, the transduced signal is modified and amplified appropriately with the help of conditioning and processing devices before passing it on to the output stage for display. Signal conditioning (by noise reduction and filtering) is performed to enhance the condition of the signal obtained in the first stage, in order to increase the signal-to-noise ratio. If required, the obtained signal is further processed by means of integration, differentiation, addition, subtraction, digitization, modulation, etc. It is important to remember here that in order to obtain an output that is analogous to the input, the characteristics of the input signals should be transformed with true fidelity.

Table 12.1 Examples of the three stages of a generalized measurement system

Type	Detector–transducer stage	Intermediate modifying stage	Terminating stage
Mechanical	Spring mass Contact spindle Bourdan tube Bellows Proving ring	Gears Cams Linkages	*Indicators*: Moving pointer and scale Light beam and scale Electron beam scale CRO, liquid column *Digital types*: Direct reading *Recorders*: Digital printing Hot stylus and chart Inked pen chart, light beam and photographic film, computer recording Magnetic recording *Controlling devices*: All types
Hydraulic-pneumatic	Orifice meter Venturi meter Vane Propeller	Valves Dash pots	
Optical	Photo-electric cell Photographic film	Lens Mirrors Light levers Filters	
Electrical	Resistance Capacitance Inductance Piezo-electric crystal Thermocouple	Amplifying systems Attenuating systems Filters Tele-metering systems	

Transducers

After studying this chapter, the reader will be able to

- understand the basic structure of several types of transducers
- elucidate the characteristics of several types of transducers
- appreciate the significance of system response
- explain the different functional elements of generalized measurement systems
- understand the inherent problems of mechanical systems
- appreciate the importance of electrical and electronic systems
- understand the working of a cathode ray oscilloscope and its importance

13.1 INTRODUCTION

We know that the generalized measuring system consists of three functional elements: primary detector-transducer stage, intermediate modifying stage, and output or terminating stage. Each stage performs certain functions so that the value of the physical variable to be measured (measurand) is displayed as output, as discussed in Chapter 12. In addition, a controlling function is required in many applications. For example, measurement systems employed in process control comprise a fourth stage called feedback control stage. The feedback control stage essentially consists of a controller that interprets the measured signal, depending on which decision is taken to control the process. Consequently, there is a change in the process parameter that affects the magnitude of the sensed variable. It is essential to note here that more the accuracy of measurement of the control variable, the better the accuracy of control. Hence, efforts have to be made towards accurate measurements before making any attempt to control the same. Figure 13.1 shows a simple schematic representation of a generalized measuring system.

13.2 TRANSFER EFFICIENCY

The detecting or sensing element of a measuring system first makes contact with the quantity to be measured, and the sensed information is immediately transduced into an analogous form. The transducer, which may be electrical, mechanical, optical, magnetic, piezoelectric, etc., converts the sensed information into a more convenient form. A transducer is a device that converts one form of energy into another form.

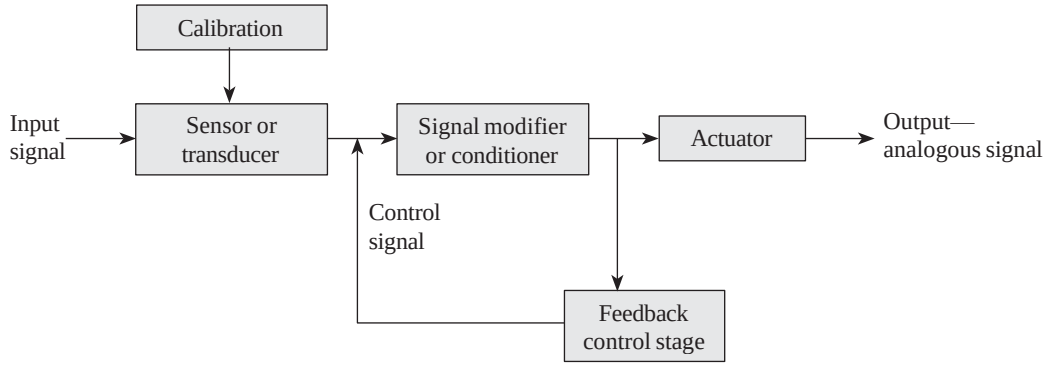


Fig. 13.1 Four stages of a generalized measuring system

The term transfer efficiency was first defined by Simpson (1955) for comparing the different devices used in the first stage. Depending on the information sensed and delivered by the sensor, transfer efficiency is defined as follows:

$$\text{Transfer efficiency} = \frac{I_{del}}{I_{sen}}$$

I_{del} is the information delivered by the pickup device and I_{sen} is the information sensed by the pickup device.

This ratio cannot be more than unity since the sensor device cannot generate any information on its own. Hence, it is desirable to have a sensor with very high transfer efficiency.

13.3 CLASSIFICATION OF TRANSDUCERS

Transducers are classified as follows:

1. Primary and secondary transducers
2. Based on the principle of transduction
3. Active and passive transducers
4. Analog and digital transducers
5. Direct and inverse transducers
6. Null and deflection transducers

13.3.1 Primary and Secondary Transducers

Transducers are generally defined as devices that transform values of physical quantities in the form of input signals into corresponding electrical output signals. The physical quantity may be heat, intensity of light, flow rate, etc. In terms of energy conversion, a transducer is defined as a device that transforms energy from one form to another. The energy may be in electrical, mechanical, or acoustical form. Transducers are devices that are used to transform the information sensed (signals) between two different physical domains. Transducers, which include the entire detector transducer or primary stage of the generalized measurement system, may comprise two important components:

Sensing or detecting element The function of this element is to respond to a physical phenomenon or a change in the physical phenomenon. Hence it is termed a primary transducer.

Transduction element The function of a transduction element is to transform the output obtained by the sensing element to an analogous electrical output. Hence it is termed a secondary transducer.

In addition to these, the primary stage also consists of auxiliary sources of energy, power amplifiers, and calibration sources.

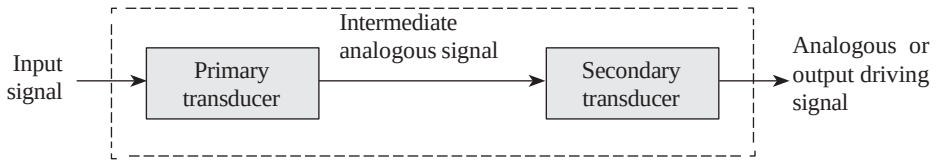


Fig.13.2 Primary detector transducer stage with primary and secondary transducers

Table 13.1 Typical mechanical primary detector transducers

Mechanical transducers	Typical operation
Contacting spindle, pin, or finger	Displacement to displacement
Elastic members Bourdon tube Bellows Diaphragm	Pressure to displacement
Proving ring Spring	Force to displacement
Mass Seismic mass	Forcing function to displacement
Pendulum scale	Force to displacement
Manometer	Pressure to displacement
Thermal Thermocouple	Temperature to electric current
Bimaterial	Temperature to displacement
Temperature stick	Temperature to phase
Hydropneumatic Static Float	Fluid level to displacement
Hydrometer	Specific gravity to displacement
Dynamic Orifice Venturi Pitot tube	Velocity to pressure
Vanes	Velocity to force
Turbines	Linear to angular velocity

An example of a primary detector transducer (see Table 13.1) stage comprising both these elements is the combination of the bourdon tube and the linear variable differential transformer (LVDT) (Figs 13.2 and 13.3). The bourdon tube, which acts as a detecting element, senses pressure and gives the output in the form of displacement. This displacement is further used to move the core of LVDT, and a voltage is obtained as output. Thus, the pressure is converted into displacement, which in turn is transduced into an analogous voltage signal. Thus, the bourdon tube acts as the primary sensing element and the LVDT as the secondary transducer.

13.3.2 Based on Principle of Transduction

This classification is based on how the input quantity is transduced into capacitance, resistance, and inductance values. They are known as capacitive, resistive, and inductive transducers. Transducers can also be piezoelectric, thermoelectric, magnetostrictive, electrokinetic, and optical.



Fig. 13.3 Example of a primary detector transducer stage

13.3.3 Active and Passive Transducers

Active transducers are of a self-generating type, wherein they develop their own voltage or current output (Fig. 13.4). They do not need any auxiliary power source to produce the output. The energy required to produce the output is derived from the physical quantity being measured. Examples of active transducers are piezoelectric crystals (used for force or acceleration measurement), tachogenerators, thermocouples, and photovoltaic cells.

Passive transducers derive the power required for transduction from an auxiliary source of power (Fig. 13.5). A part of the power required for generating the output is derived from the physical quantity being measured. Since they receive power from an auxiliary source, they are termed externally powered transducers. Further, these transducers remain passive in the absence of an external power source; hence, they are called passive transducers. Resistive, capacitive, and inductive transducers are some of the examples of passive transducers.

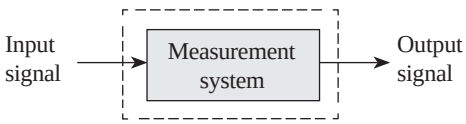


Fig. 13.4 Active transducer

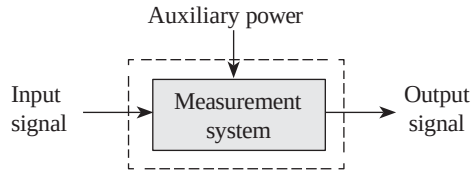


Fig. 13.5 Passive transducer

13.3.4 Analog and Digital Transducers

Based on the output generated, that is, depending on whether the output generated is a continuous function of time or is in discrete form, transducers are classified as analog and digital transducers.

In case of analog transducers, the input quantity is converted into an analog output, which is a continuous function of time. LVDT, strain gauge, thermocouple, and thermistor are some examples of analog transducers.

If a transducer converts the input quantity into an electrical signal that is in the form of pulses, as output, it is called a digital transducer. These pulses are not continuous functions of time but are discrete in nature. In fact, it is easier to process and transmit data as a function of two numbers 0 and 1. Examples are shaft encoders, linear displacement transducers using conducting or non-conducting contacts, and opaque or translucent segments.

13.3.5 Direct and Inverse Transducers

When a measuring device measures and transforms a non-electrical variable into an electrical variable, it is called a direct transducer (Fig. 13.6). A thermocouple that is used to measure temperature, radiation, and heat flow is an example of a transducer.

If an electrical quantity is transformed into a non-electrical quantity, it is termed an inverse transducer (Fig. 13.7). A very common example of an inverse transducer is a piezoelectric

crystal wherein a voltage is given as the input. This is because when a voltage is applied across its surfaces, its dimensions are changed causing a mechanical displacement.

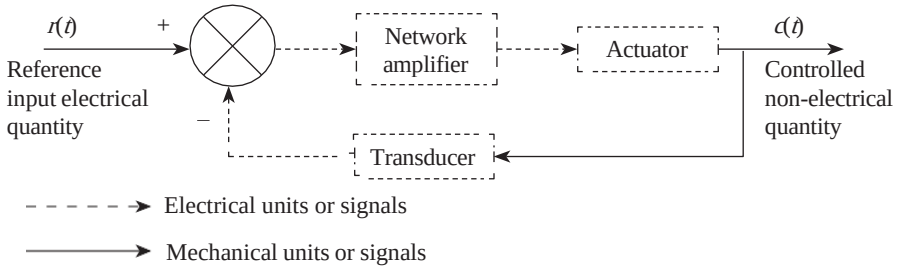


Fig. 13.6 Transducers

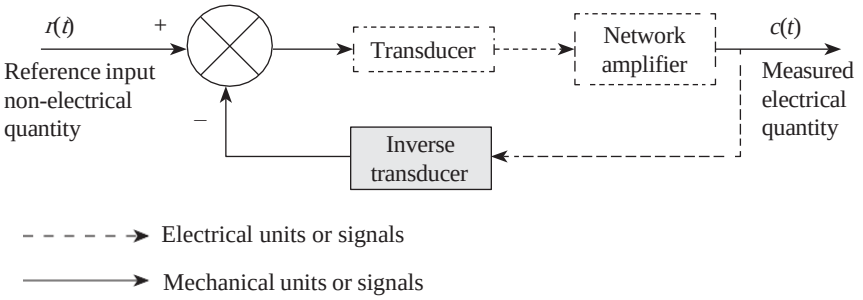


Fig. 13.7 Inverse transducers

13.3.6 Null- and Deflection-type Transducers

This classification is based on null or deflection principle. A null-type device works on the principle of maintaining a zero deflection by applying an appropriate known effect that opposes the one generated by the measured quantity. This type of transducer requires a detector capable of detecting an imbalance and a means to restore a balance either manually or automatically. (A null detector and a means of restoring the balance are necessary.) Since it is required to balance the deflection preferably at the zero level, in order to determine the numerical values, precise information of the magnitude of the opposing effect is essential. The major disadvantage of this method is that it cannot be used for dynamic measurements. Examples of null-type transducers include dead-weight pressure gauges and equal arm balances. In a dead-weight pressure gauge, which operates on the null method, standard weights are added to the platform of known weight and gravitational force can be used to balance the pressure force on the face of the piston. The balance of forces is indicated by the platform remaining stationary in between the two stops. The pressure required for the balance may be calculated as the weights added and the area of the piston are known.

A deflection-type transducer works on the principle that the measured quantity produces a physical effect that stimulates an analogous effect in the opposite direction in some parts of the measuring instrument. There is an increase in the opposing effect until a balance is restored, at which stage the measurement of deflection is carried out. A rudimentary pressure gauge is an example of a deflection transducer. It contains a piston that is driven by the fluid pressure to

produce a resultant force. The piston acts as a primary transducer and a secondary transducer, which also acts as the variable conversion element. The force generated by the piston rod stimulates an opposing force in the spring due to an imbalance of forces on the piston rod. This causes a deflection of the spring. As the spring deflection increases, its force increases and thus a balance is achieved. Another example is a spring balance system. Calibration of the spring determines the accuracy in a deflection-type instrument.

13.4 QUALITY ATTRIBUTES FOR TRANSDUCERS

A transducer should possess the following qualities:

1. *Repeatability*: A transducer must have a high degree of accuracy and repeatability.
2. *Linearity*: It should have a very high degree of linearity within the specified operating range.
3. *Dynamic response*: The dynamic response of the transducer should be instantaneous.
4. *Impedance*: It should have a high input impedance and a low output impedance for loading effect elimination.
5. *High resolution*: It should have good resolution over the complete selected range.
6. *Size*: The size of the transducer should be preferably small.
7. *Hysteresis*: The transducer should possess low or no hysteresis.
8. *Robustness*: It should be able to endure pressure, vibrations, shocks, and rough handling.
9. *Adaptability*: The transducer should be capable of working in a corrosive environment.
10. *Sensitivity*: Cross-sensitivity of the transducer should be zero or minimum, and it should exhibit high sensitivity to the desired signal.
11. *Stability and reliability*: The degree of stability and reliability of the transducer should be high.
12. *Response*: It should have a requisite time domain specification of transient and frequency response.

13.5 INTERMEDIATE MODIFYING DEVICES

The primary detector transducer stage usually detects a mechanical quantity and transduces it into an electrical form. In the intermediate modifying stage of the measurement system, the transduced signal is further modified and amplified with the help of conditioning and processing devices, so that the signal is appropriate before passing it on to the output or terminating stage for display. The terminating-stage elements usually comprise indicators, recorders, display and data processing units, and control elements.

Signal conditioning is essential in order to enhance the condition of the signal obtained in the first stage by reducing the noise and filtering such that the signal-to-noise ratio is increased. If required, the obtained signal has to be processed further by means of attenuation, integration, differentiation, addition, subtraction, digitization, modulation, etc. It is important to remember here that, in order to obtain an output that is analogous to the input, the characteristics of the input signals should be transformed for display with true fidelity. Hence, the dynamic measurement of mechanical quantities imposes severe strain on the elements of the intermediate stage.

Table13.2 Typical electrical, electronic, and self-generating primary detector transducers

Electrical transducers	Typical operation
Resistance Contacting Variable area of conductor Variable length of conductor	Displacement to change in resistance
Variable dimensions of conductor	Strain to change in resistance
Variable resistivity of conductor	Air velocity to change in resistance
Inductance Variable coil dimensions Variable air gap Changing core material Changing coil positions Change core positions	Displacement to change in inductance
Moving coil Moving permanent magnet Moving core	Velocity to change in inductance
Capacitance Changing air gap Changing plate areas Changing dielectric	Displacement to change in capacitance
Self-generating Piezoelectric pickup Photovoltaic Moving coil generator Thermocouple	Displacement to voltage Intensity of light source to voltage Displacement to voltage Temperature or radiation to voltage
Electronic	Displacement to current
Photoelectric Photoconductive Photogenerative Photoemissive	Intensity of light source to voltage
Electrokinetic	Flow to voltage

One of the major requirements of intermediate modifying devices is large amplification coupled with good transient response. Mechanical, hydraulic, and pneumatic devices fulfil both these requirements inadequately and hence electrical and electronic devices are preferred (Table 13.2).

13.5.1 Inherent Problems in Mechanical Systems

We know that primary detector transducers transduce the input signal into mechanical displacement. This intermediate analogous signal is passed on to the secondary transducer and gets converted into an electrical form, which is more easily processed by the intermediate stage.

CHAPTER-4

Measurement of Force, Torque, and Strain

After studying this chapter, the reader will be able to

- understand the basics and describe methods of force, torque, and strain measurements
- explain the direct method of force measurement
- differentiate between the different types of elastic members for force measurements
- discuss the different electrical pressure transducers
- elucidate the various types of dynamometers used for torque measurements
- appreciate the different aspects of strain measurement systems
- explain gauge factor measurement
- describe temperature compensation

14.1 INTRODUCTION

We all know that force is defined as the product of mass and acceleration, as per Newton's second law of motion. Force is a vector quantity whose unit of measurement is Newton (N); 1 N is equivalent to the force required to accelerate a mass of 1 kg at a rate of 1 m/s^2, that is, 1 N = 1 kg × 1 m/s^2. The standards of mass and acceleration determine the standard of force. It is a well-known fact that the fundamental unit of mass is kilogram, which is traceable to the international prototype of mass. The international standard for mass is international kilogram, which is equivalent to the mass of a cylinder of platinum–iridium maintained at Sevres, France. Although acceleration is not a fundamental unit, it is derived from two other well-defined and internationally accepted fundamental quantities: length and time. The acceleration due to gravity g can be adopted as a suitable standard and is approximately equal to 9.81 m/s^2.

Torque is defined as the force, which, besides exerting effects along its line of action, may also exert a rotating force relative to any axis excepting those intersecting the line of action. It can also be termed as a couple or twisting force. It is determined by measuring the force at a known radius r and is expressed as $T = F \times r$. The metric unit of measurement of torque is Newton metre (N m), which is a force multiplied by the distance of this force from an axis of rotation. Torque is closely related to the measurement of force, and since force and length that define torque conform to international standards, separate standards are not required for torque.

14.2 MEASUREMENT OF FORCE

The methods for measuring force can be classified into two basic categories: direct and indirect. In case of direct methods, a direct comparison is made between an unknown force and the known gravitational force on a standard mass. For this purpose, a beam balance may be employed wherein masses are compared. In this case, the beam neither attenuates nor amplifies. Indirect comparison is made by a calibrated transducer that senses gravitational attraction or weight. Sometimes, the deformation due to a force applied on an elastic member is measured.

14.2.1 Direct Methods

Direct methods involve the comparison of an unknown force with a known gravitational force on the standard mass. A force is exerted on a body of mass m due to the earth's gravitational field, which can be represented by the following equation:

$$W = mg$$

Here m is the standard mass, g is the acceleration due to gravity, and W is the weight of the body.

It is imperative to know the values of mass and acceleration due to gravity accurately in order to determine the force acting on the body. With the help of an analytical balance, a direct comparison can be drawn between an unknown force and the gravitational force. The following section describes the working of an analytical balance.

Analytical Balance

An analytical balance, also known as an equal arm balance, is probably the simplest force-measuring system. As discussed in Section 14.2, an unknown force is directly compared with a known gravitational force. Comparison of masses is carried out by attaining some kind of beam balance by employing a null balance method. It is sufficient to only find out the magnitude since the unknown force and the gravitational force act in directions parallel to each other. The working principle is illustrated schematically in Fig. 14.1.

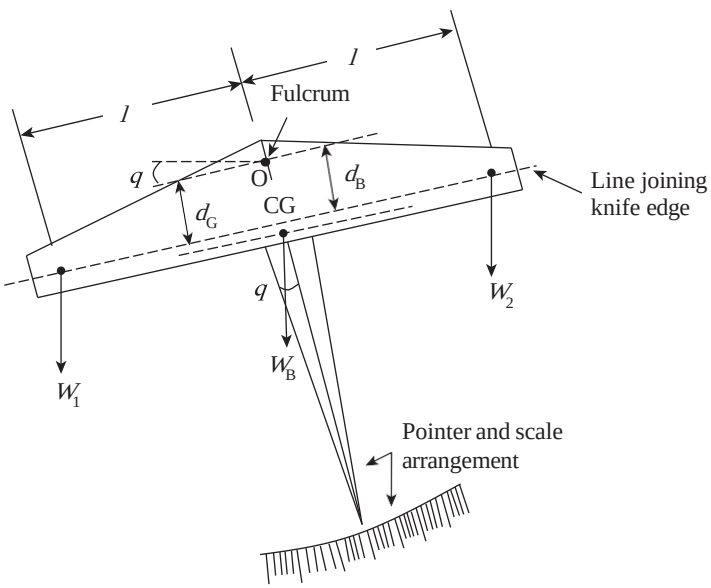


Fig. 14.1 Analytical balance

The rotation of the balance arm is about the knife edge point or fulcrum O. Figure 14.1 shows the balance in an unbalanced condition. The distance between the fulcrum and the centre of gravity point CG is d_G. Let W_B denote the weight of the balance arms and pointer; W_1 and W_2 are two weights acting on either side of the balance. When the two weights W_1 and W_2 are equal, angle q will be zero, and hence the weight of the balance arms and pointer will

not influence the measurements.

Deflection per unit unbalance gives a measure of the sensitivity of the balance. The difference between the two weights, that is, $W_1 - W_2$, gives the unbalance. Let this difference be ΔW.

$$\text{Hence, sensitivity } S = \frac{q}{(W_1 - W_2)} = \frac{q}{\Delta W} \tag{14.1}$$

When the balance is at equilibrium, the following equation is realized:

$$W_1(l \cos q - d_B \sin q) = W_2 (l \cos q + d_B \sin q) + W_B\, d_G \sin q \tag{14.2}$$

When the angle of deflection is small, $\sin q = q$ and $\cos q = 1$. For such small angles of deflection, Eq. (14.1) can be modified as follows:

$$W_1 (l - d_B q) = W_2 (l + d_B q) + W_B\, d_G\, q$$

$$\text{Therefore, } q = \frac{l\,(W_1 - W_2)}{(W_1 + W_2)\, d_B + W_B\, d_G} \tag{14.3}$$

$$\text{Hence, sensitivity } S = \frac{q}{(W_1 - W_2)} = \frac{l}{(W_1 + W_2)\, d_B + W_B\, d_G} \tag{14.4}$$

When in equilibrium, $W_1 = W_2 = W$.

$$S = \frac{l}{2W d_B + W_B\, d_G} \tag{14.5}$$

If balances are constructed such that d_B, the distance between the fulcrum and the line joining the knife edges, is zero, then equation (14.5) becomes as follows:

$$S = \frac{l}{W_B\, d_G} \tag{14.6}$$

Hence, it can clearly be seen from Eq. (14.6) that the sensitivity is independent of the applied weights W_1 and W_2. There may be errors in measurement due to the buoyancy effects of air exerted on the weights, which can be eliminated through proper care. In addition, to enhance sensitivity, it is necessary to increase l and decrease W_B and d_G, which again affect the stability. Hence, a trade-off needs to be made between the stability and sensitivity of the balance. It is pertinent to mention here that for optimal performance, the design and operation of the balance is crucial. The disadvantage associated with an analytical balance is that it necessitates the use of a set of weights that are at least equal to the maximum weight to be measured. However, it can be mentioned here that an equal-arm balance system is not appropriate for the measurement of large weights.

Platform Balance

For applications involving measurement of large weights, a platform balance or multiple-lever system is preferred. A platform balance is schematically represented in Fig. 14.2.

In a platform balance, two smaller weights, poise weight W_x and pan weight W_y, are used for the measurement of a large weight W. The initial zero balance can be set by an adjustable

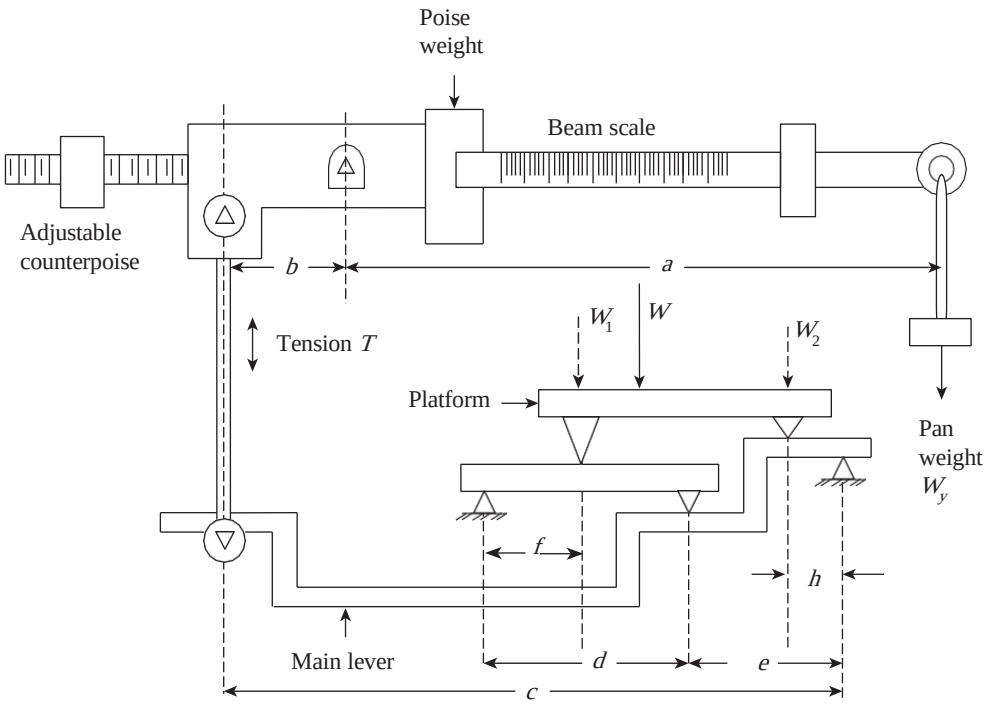


Fig. 14.2 Platform balance

counterpoise. The poise weight W_x is set to zero of the beam scale, and then the counterpoise is adjusted for obtaining the initial balance before applying the unknown weight W on the platform. For ease of analysis, it can be assumed here that the two weights W_1 and W_2 are substituted by W. Since W_x has already been adjusted to zero, the entire unknown weight W is balanced by the pan weight W_y.

It can be seen that $T \times b = W_y \times a$ (14.7)

and $T \times c = W_1(f/d)e + W_2 \times h$ (14.8)

The linkage proportion can be arranged such that $h/e = f/d$.

Hence, we have $T \times c = (W_1 + W_2)h = Wh$ (14.9)

It is clear from these discussions that weight W can be placed anywhere on the platform and its position with respect to the knife edges of the platform is not significant.

From Eqs (14.7) and (14.9), we obtain the following equation:

$$\frac{W_y a}{b} = \frac{Wh}{c} \tag{14.10}$$

This gives a new equation.

$$W = \frac{a}{b}\frac{c}{h} W_y = S\, W_y \tag{14.11}$$

where S is known as the scale multiplication ratio, which is given by another equation:

$$S = \frac{a}{b}\frac{c}{h} \tag{14.12}$$

The multiplication ratio gives an indication of the weight that should be applied to the pan in order to balance the weight on the platform. This means that if the multiplication factor is 100, then a weight of 1 kg applied on the pan can balance a weight of 100 kg placed on the platform. Let us assume that the beam is divided with a scale of m kg per scale division, then a poise movement on the y scale division has to produce the same result as a weight W_x placed on the pan at the end of the beam; hence, $W_x y = mya$. Therefore, we obtain the following equation:

$$m = W_x/a \tag{14.13}$$

The required scale divisions on the beam for any poise weight W_x is determined by this relationship. The length of the beam scale a is expressed in terms of scale divisions. The beam is balanced by appropriate combinations of pan weights and an adjustment of the poise weight along the calibrated beam scale.

14.3 ELASTIC MEMBERS

For determining the applied force, many force-measuring transducers employ various mechanical elastic members or their combinations. Application of a load or force on these members causes an analogous deflection. This deflection, which is usually linear, is measured either directly or indirectly by employing secondary transducers. This displacement is then converted into an electrical signal by the secondary transducers. The strain gauge is the most popular secondary transducer employed for the measurement of force. Figure 14.3 depicts the use of a strain gauge for the determination of applied force.

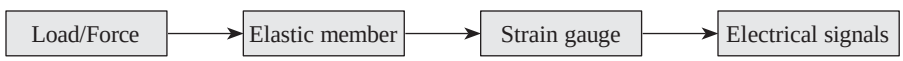


Fig. 14.3 Use of a strain gauge for force determination

The strain gauge operates on a simple principle. When force is applied on an elastic member, a steel cylindrical rod in this case, the dimensions undergo a change. If the strain gauge is bonded to the cylinder, the gauge is stretched or compressed, resulting in a change in its length and diameter. The dimension of the strain gauge changes due to a change in resistance. This change in resistance or output voltage of the strain gauge gives a measure of the applied force. Load cells and proving rings that employ strain gauges are the two most common instruments used for force measurement.

14.3.1 Load Cells

Elastic members are generally used for the measurement of force through displacement measurement. When an elastic member is combined with a strain gauge and used for the measurement of force, it is termed as a load cell. In load cells, elastic members act as primary transducers and strain gauges as secondary transducers. A load cell is used in an indirect method of force measurement where force or weight is converted into an electrical signal. Load cells are used extensively for the measurement of force.

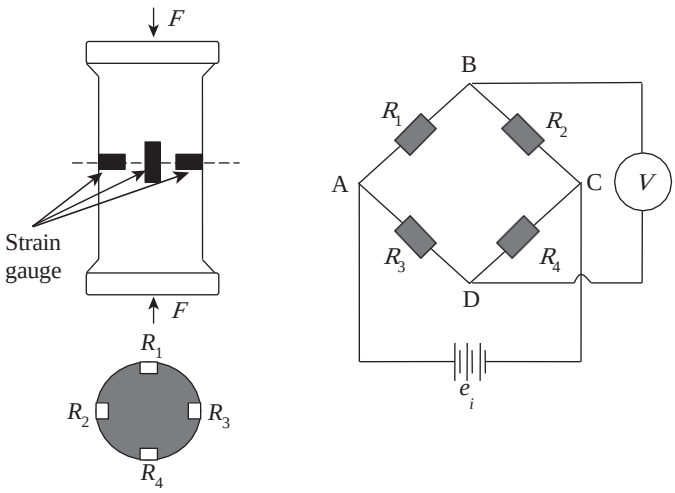


Fig. 14.4 Strain gauge arrangement for a load cell

A load cell comprises four strain gauges; two of these are used for measuring the longitudinal strain while the other two for measuring the transverse strain. The four strain gauges are mounted at 90° to each other, as shown in Fig. 14.4.

Two gauges experience tensile stresses while the other two are subjected to compressive stresses. At the no-load condition, resistance in all the four gauges will be same. The potential across the two terminals B and D are same. The Wheatstone bridge is now balanced and hence output voltage is zero When the specimen is stressed due to the applied force, the strain induced is measured by the gauges. Gauges R_1 and R_4 measure the longitudinal (compressive) strain, while gauges R_2 and R_3 measure the transverse (tensile) strain. In this case, voltages across the terminals B and D will be different, causing the output voltage to vary, which becomes a measure of the applied force upon calibration.

The longitudinal strain developed in the load cell is compressive in nature and given by the following relation:

$$e_1 = -\frac{F}{AE}$$

Here, F is the force applied, A is the cross-sectional area, and E is the Young's modulus of elasticity.

Gauges 1 and 4 experience this strain. Strain gauges 2 and 3 would experience a strain given by the following relation:

$$e_2 = -\frac{g\,F}{AE}$$

Here, g is the Poisson's ratio.

This arrangement of mounting of gauges compensates the influence of bending and temperature. In fact, a complete compensation is obtained if the gauges are mounted symmetrically.

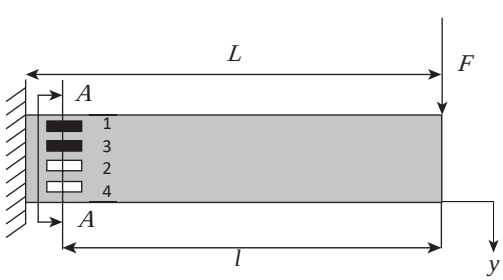


Fig. 14.5 Cantilever beam for force measurement

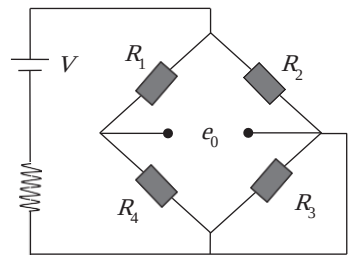


Fig. 14.6 Strain gauge arrangement for force measurement

14.3.2 Cantilever Beams

A cantilever beam is one of the simple devices used for the measurement of force. One end of the cantilever is fixed and a force F is applied on the free end, as shown in Fig. 14.5. Four strain gauges are bonded near the fixed end to measure the strain induced by the application of force. Two strain gauges are fixed on the top for measuring tensile strain $+e$ and two at the bottom side for measuring compressive strain $-e$. The strain developed at the fixed end is determined using the following equation:

$$e = \frac{6Fl}{Ebt^2}$$

Here, l is the length of the beam, E is the Young's modulus of the material, b is the width of the beam, and t is the thickness of the cantilever.

In Fig. 14.6, strain gauges R_1 and R_3 measure tensile strain, whereas compressive strain is measured by R_2 and R_4. Strains measured by all the four gauges are of equal magnitude. Apart from the measurement of force, a load cell finds application in vehicle weighbridges and cutting tool force dynamometers.

14.3.3 Proving Rings

One of the most popular devices used for force measurement is the proving ring. In order to measure the displacement caused by the applied pressure, a displacement transducer is connected between the top and bottom of the ring. Measurement of the relative displacement gives a measure of the applied force. A proving ring can be employed for measuring the applied load/force, with deflection being measured using a precise micrometer, a linear variable differential transformer (LVDT), or a strain gauge. When compared to other devices, a proving ring develops more strain owing to its construction. A proving ring, which is made up of steel, can be used for measuring static loads, and is hence employed in the calibration of tensile testing machines. It can be employed over a wide range of loads (1.5 kN to 2 MN).

A proving ring comprises a circular ring having a rectangular cross-section, as shown in Fig. 14.7. It has a radius R, thickness t, and an axial width b. The proving ring may be subjected to either tensile or compressive forces across its diameters. The two ends between which force is measured are attached with structures. Four strain gauges are mounted on the walls of the proving ring, two on the inner walls, and two other on the outer walls. The applied force induces a strain (compressive) $-e$ in gauges 2 and 4, while gauges 1 and 3 undergo tension

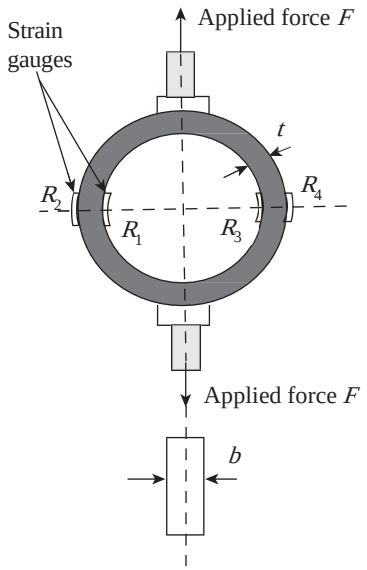


Fig. 14.7 Proving ring

+*e*. The four strain gauges are connected to the bridge circuit, and the unbalanced voltage caused by the applied force can be measured. This voltage is calibrated in terms of force. The magnitude of the strain is given by the following expression:

$$e = \frac{1.08FR}{Ebt^2}$$

The relationship between the applied force and the deflection caused by the applied force is given by

$$d_y = \left(\frac{2\pi - 4}{\pi^4}\right)\frac{16Fd^3}{EI}$$

Here, E is the Young's modulus, I is the moment of inertia, F is the force, d is the outside diameter of the ring, and d_y is the deflection.

14.3.4 Differential Transformers

An LVDT can be employed in a load cell to measure the applied force. The working of an LVDT has been discussed in detail in Chapter 16 (refer to Section 16.7.3). An LVDT can be used in conjunction with a primary transducer to measure force, pressure, weight, etc. It can generate an AC output voltage as a function of the displacement of the magnetic core caused by the applied force. The frictionless movement and low mass of the core provide advantages such as high resolution and low hysteresis, making the LVDT an obvious choice for dynamic measurement. LVDTs can be employed as alternatives to strain gauges or micrometers.

14.4 MEASUREMENT OF TORQUE

Measurement of torque is important for the following reasons:

1. It is necessary to obtain load information for the analysis of stress or deflection. In this case, torque T is determined by the measurement force F at a known radius r using the following relationship:

   $T = Fr$ (in N m)

2. Measurement of torque is essential in the determination of mechanical power. Mechanical power is nothing but the power required to operate a machine or the power developed by the machine, and is computed using the following relation:

   $P = 2\pi NT$

   Here, N is the angular speed in resolution per second.

   Torque-measuring devices employed for this purpose are popularly known as dynamometers. Dynamometers are used to measure torque in internal combustion machines, small steam turbines, pumps, compressors, etc.

3. Measurement of torque is important for evaluating the performance characteristics of machines.

The nature of the machine to be tested determines the kind of dynamometer to be used for torque measurement. If the machine generates power, the dynamometer employed must

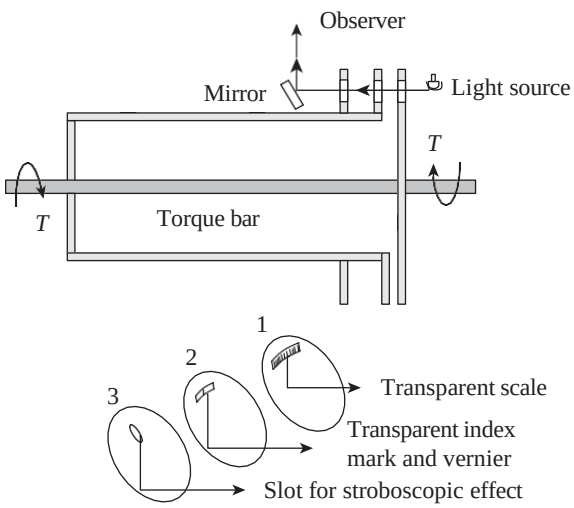


Fig. 14.8 Torsion-bar dynamometer

have the capability of absorbing that power. These dynamometers are called absorption dynamometers, which are particularly useful for measuring the power or torque developed by power sources such as engines or electric motors.

In case the machine is a power absorber, the dynamometer must have the capacity to drive it; such a dynamometer is known as a driving dynamometer. Driving dynamometers are therefore useful in determining performance characteristics of devices such as pumps and compressors. The third type is the transmission dynamometer. These are passive devices placed at an appropriate location within the machine or between machines. They are used to sense torque at that particular location. These are sometimes known as torque meters.

14.4.1 Torsion-bar Dynamometer

Elastic deflection of the transmitting element may be used for the measurement of torque, which can be achieved by measuring either a gross motion or a unit strain. The main problem associated with either case is the difficulty in reading the deflection of the rotating shaft. A torsion-bar dynamometer is also known as a torsion-bar torque meter, which employs optical methods for deflection measurement, as shown in Fig. 14.8.

Calibrated scales are used to read the relative angular displacement of the two sections of the torsion bar. This is possible because of the stroboscopic effect of intermittent viewing and persistence of vision. Transmission dynamometers, which employ this principle, are available in ranges up to 60,000 m kgf and 50,000 r/min, having an error of ±0.25%.

Replacing the scales on disks 1 and 2 with sectored disks, which are alternately transparent and opaque sectors, and the human eye with an electro-optical transducer, a version having an electrical output is obtained. When there is no torque, the sectored disks are positioned to give a 50% light transmission area. The area of proportionality increases with positive torque and decreases with negative torque, thus giving a linear and direction-sensitive electric output.

14.4.2 Servo-controlled Dynamometer

Torque and speed are measured under actual driving conditions of an automobile engine; tape recordings of such an exercise of an engine are obtained and then simulated under laboratory conditions. Engine speed and torque are controlled by two feedback systems. A servo-controlled dynamometer is schematically represented in Fig. 14.9.

The actual speed signal generated by the tachometer generator, from the dynamometer, is compared with the preferred speed that is set in the tape recorder. If the actual and the preferred speeds are not the same, the dynamometer control is automatically adjusted until they are equal. The load cell on the dynamometer measures the actual torque from the engine, and is compared with the preferred torque that is set in the tape recorder. If these two values

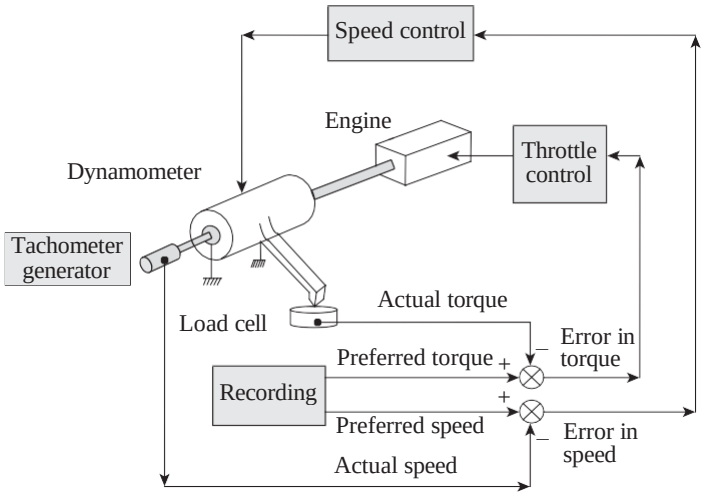


Fig. 14.9 Schematic representation of a servo-controlled dynamometer

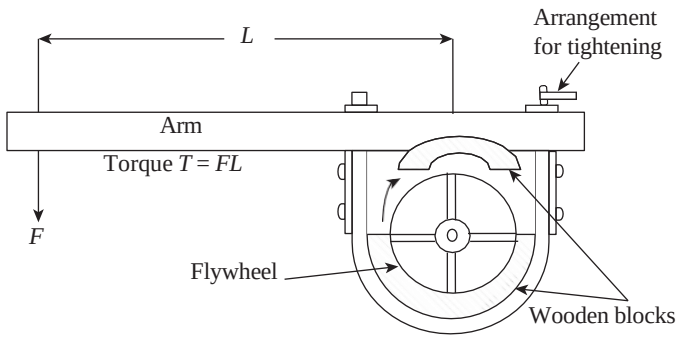


Fig. 14.10 Prony brake dynamometer

differ, then the error signal generated actuates the engine throttle control in the appropriate direction. Both torque control and speed control operate simultaneously and continuously such that they conform to the desired value set in the tape recorder.

14.4.3 Absorption Dynamometer

The Prony brake dynamometer was invented by Gaspard de Prony, a French engineer inventor, in 1821 to measure engine power. A Prony brake dynamometer is one of the simplest, inexpensive, and the most popular absorption dynamometers. It is a mechanical type of device that depends on dry friction wherein mechanical energy is converted into heat.

As illustrated in Fig. 14.10, a Prony brake dynamometer comprises two wooden blocks that are mounted on either side of the fly wheel in diagrammatically opposite directions. The fly wheel is attached to the shaft whose power needs to be determined. A lever arm is fixed to one block and the other arm is connected to an arrangement provided to tighten the rope. Tightening of the rope is performed in order to enhance the frictional resistance between the blocks and the flywheel.

The torque exerted by the Prony brake is given by the following equation:

$$T = FL$$

Here, force F is measured by conventional force-measuring instruments such as load cells or balances.

The power dissipated in the brake is then calculated by the following equation:

$$P = \frac{2\pi NT}{60}$$

or

$$P = \frac{2\pi NFL}{60} W$$

Here, P is the dissipated power in watts, L is the length of the lever arm in metres, N is the angular speed in revolution per minute, and F is the force in Newton.

Although a Prony brake is inexpensive, it is inherently unstable. It is difficult to adjust or

maintain a specific load. The following are certain limitations associated with a Prony brake dynamometer:

1. Due to wear of the wooden blocks, there will be variations in the coefficients of friction between the blocks and the flywheel. This necessitates tightening of the clamp. This makes the system unstable, and large powers cannot be measured particularly when used for longer periods.
2. The coefficients of friction decrease due to excessive rise in temperature, which may result in brake failure. Therefore, cooling is required in order to limit the temperature rise. Water is supplied into the hollow channel of the flywheel to provide cooling.
3. Due to variations in coefficients of friction, there may be some difficulty in taking readings of force F. The measuring arrangement may be subjected to oscillations, especially when the machine torque is not constant.

14.5 MEASUREMENT OF STRAIN

The problem of determining stresses acting on a body has been an important and interesting aspect of materials engineering. Prior to the invention of electrical resistance, strain gauge extensometers were extensively employed for the determination of strain. These were associated with a number of problems. One of the most important problems was their bulk, which was a major constraint in its use. Due to lack of precise knowledge of stress conditions of the material or structure, many engineering problems were solved on a theoretical basis, by employing trial and error methods and by assuming a large factor of safety.

In 1856, Lord Kelvin established the principle of electrical resistance strain gauges when he demonstrated that when stress is applied to a metal wire, apart from undergoing changes in both length and diameter, there will be some changes in the electrical resistance of the wire.

When external forces act on a body, the body is in a condition of stress and strain. Since direct measurement of stress is not possible, its effects, such as change of shape of the body or change in length, are measureable, thus providing a known relationship between stress and strain. If adequate information is available, stresses acting on a body can be computed.

When a force or load is applied to a body, it undergoes some deformation. This deformation per unit length is known as unit strain or simply strain, denoted by e, and is given by the following equation:

$$e = \frac{\delta l}{l}$$

Here, δl is the change in the length of the body and l is the original length of the body.

Strain gauges are essentially employed for two different purposes:

1. To determine the state of strain existing at a point on a loaded member for carrying out strain analysis
2. To act as a strain-sensitive transducer element in the measurement of quantities such as force, pressure, displacement, and acceleration

Normally, measurements are made over the shortest gauge lengths. The change in length measured over a finite length does not give the value of strain at a fixed point but rather gives the average strain over the entire length. A magnification system is essential since the change in

length is over a small gauge length. Two types of strain gauges are employed:
1. Mechanical strain gauges
2. Electrical strain gauges

14.5.1 Mechanical Strain Gauges

The Berry strain gauge extensometer, shown in Fig. 14.11, is used in civil engineering for structural applications. This gauge was designed by Professor Herman C. Berry of the University of Pennsylvania in 1910. It is employed for measuring small deformations under linear strain conditions over gauge lengths up to 200 mm.

The strain obtained is magnified by employing mechanical means. Lever systems are employed for amplifying the displacements measured over the gauge lengths. Mechanical strain gauges comprise two gauge points: one is a fixed point and the other is connected to the magnifying lever. Both points are placed on the specimen. The displacement caused is magnified by the lever and is indicated in the dial indicator. Strain is determined by dividing the measured displacements over the gauge length.

Earlier extensometers used a single mechanical lever for obtaining a magnification of 10 to 1. This was sufficient to work on a long gauge length. Recent extensometers, which are used for short gauge length, employ compound levers (dial gauges) having a higher magnification of 2000 to 1.

One of the main advantages of a mechanical strain gauge is that it contains a self-contained magnification system. There is no need to use auxiliary equipment in a mechanical strain gauge. It is best suited for conducting static tests. The following are some obvious disadvantages associated with a mechanical strain gauge:
1. Its response is slow due to high inertia, and friction is more.
2. Automatic recording of the readings is not possible.
3. It cannot be used for dynamic strain measurements and varying strains.

14.5.2 Electrical Strain Gauges

In electrical strain gauges, a change in strain produces a change in some electrical characteristic

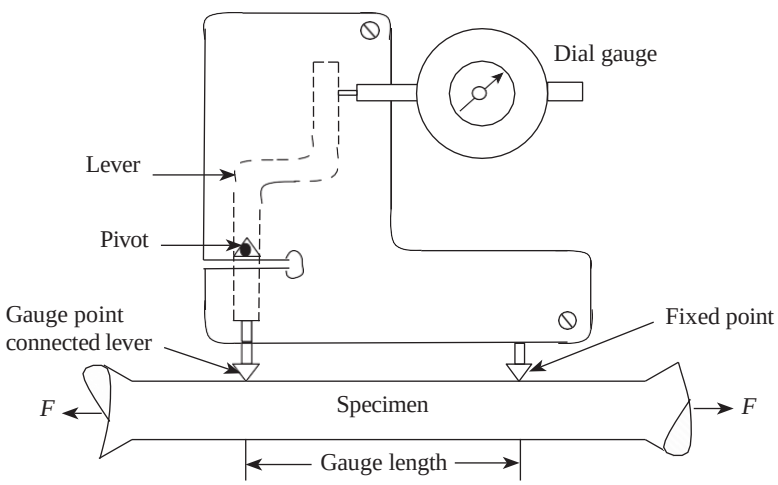


Fig. 14.11 Berry strain gauge extensometer

such as capacitance, inductance, or resistance. Magnification of the measured strain can be attained by employing an auxiliary equipment. The main advantages of measurement of strain using electrical methods are their high sensitivity and ability to respond to dynamic strains. The following are the advantages of these strain gauges:

1. They are simple in construction.
2. Strain gauges can be calibrated in terms of quantities such as force, displacement, pressure, and acceleration.
3. They are very sensitive.
4. Linear measurement is accomplished.
5. They provide good output for indicating and recording purposes.
6. Strain gauges are inexpensive and reliable.

There are two types of resistance strain gauges: unbonded and bonded.

Unbonded Resistance Strain Gauges

In case of unbonded resistance strain gauges, the electrical resistance element or the grid is not supported. A fine wire is stretched out between two or three points, which may become part of a rigid base that itself is strained, as shown in Fig. 14.12.

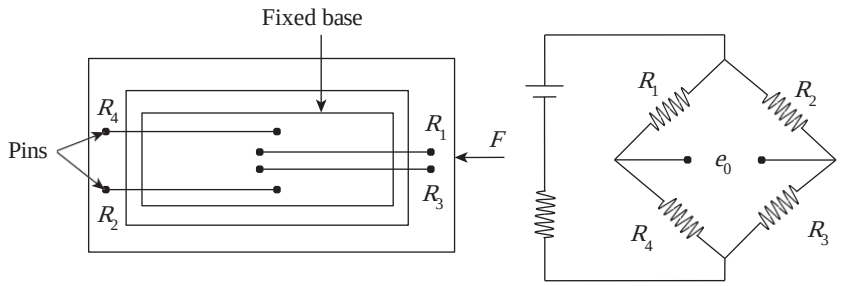


Fig. 14.12 Unbonded resistance strain gauge

The resistance strain gauge shown in Fig. 14.12 can be used to measure compressive strain as it is preloaded. The four resistance wires are connected in such a manner that they act as a full bridge. Unbonded strain gauges are used more often as elements of force and pressure transducers, and also in accelerometers. They are seldom used for strain measurement.

Bonded Resistance Strain Gauges

When a strain gauge, which is an appropriately shaped piece of resistance material, is bonded to the surface of the material to be tested, it is known as a bonded strain gauge. As shown in Fig. 14.13, a thin wire in the form of a grid pattern is cemented in between thin sheets of

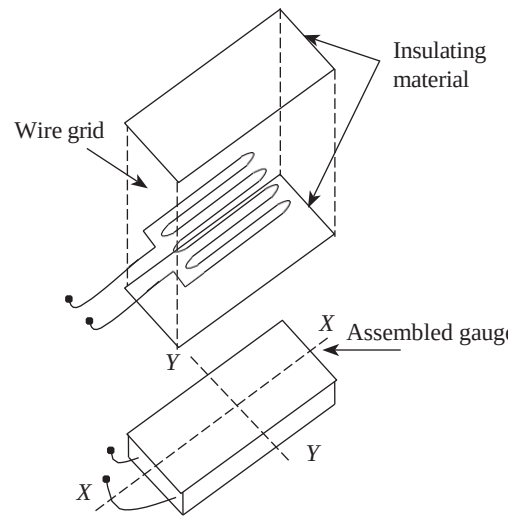


Fig. 14.13 Exploded view of a bonded resistance strain gauge

insulating materials such as paper or plastic. The strain gauge may also be a thin metal foil. The gauges are bonded to the surface under study using a thin layer of adhesive, and waterproofing is provided by applying a layer of wax or lacquer on the gauge. The strain experienced by the grid and the surface under study to which the grid is bonded will be the same. The gauge is highly sensitive in the axial direction, and in the transverse direction strain occurs due to the Poisson's ratio effect; hence, some small error may be present in the measurement of strain. However, cross-sensitivity can be completely eliminated by making the ends of the foil gauges thick. A cluster of three or more strain gauges, known as rosette gauges, can be employed if the direction of the principal strain is unknown.

Rosette gauges are normally used in the measurement of strain in complex parts. These gauges comprise three or four separate grids having different orientations and can be cemented to the surface of the part. The resultant strain can then be gathered and recorded. From the data gathered, the true magnitudes and directions of the significant surface strains can be computed.

The resistance type of electric strain gauges is the most popular device used for measuring strain. In this case, electrical resistance of the gauge proportionately changes with the strain it is subjected to. These gauges exist in the following forms:

1. Wire-type strain gauge
2. Foil-type strain gauge
3. Semiconductor or piezoresistive strain gauge

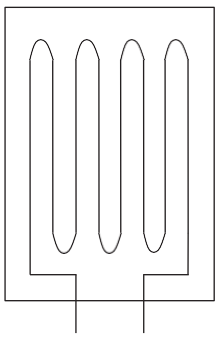


Fig. 14.14 Wire-type strain gauge

Wire-type strain gauge It is made up of a very thin wire having a diameter of 0.025 mm wound into a grid shape, as illustrated in Fig. 14.14. The grid is cemented in between two pieces of thin paper with ceramic or plastic backing. This will facilitate easy handling of the strain gauge. This gauge is further bonded securely to the surface of the member on which the strain is to be measured using a suitable cement.

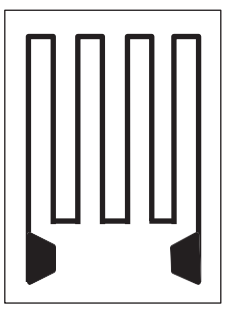


Fig. 14.15 Foil-type resistance strain gauge

Foil-type resistance strain gauge A foil-type resistance strain gauge has a thickness of about 0.005 mm. It comprises a metal foil grid element on an epoxy support. For high-temperature applications, epoxy filled with fibre glass is employed. These gauges are manufactured by printing them on a thin sheet of metal alloy with an acid-resistant ink and the unprinted portion is etched away. They offer advantages such as improved hysteresis, better fatigue life, and lateral strain sensitivity. They can be fixed on fillets and sharply curved surfaces because they are thin and have higher flexibility. These are also known as metallic gauges. A foil-type resistance strain gauge is shown in Fig. 14.15.

Semiconductor or piezoresistive strain gauge Semiconductor gauges are cut from single crystals of silicon or germanium. This type of gauge contains precise amounts of impurities like boron, which impart certain desirable characteristics. The backing, bonding materials, and mounting techniques used for metallic gauges are also employed here. When the material onto which semiconductor gauges are bonded is strained, the current in the semiconductor material changes correspondingly. When the electrical resistance of the gauge increases due to the tensile strain, it is known as a positive or p-type semiconductor gauge. When

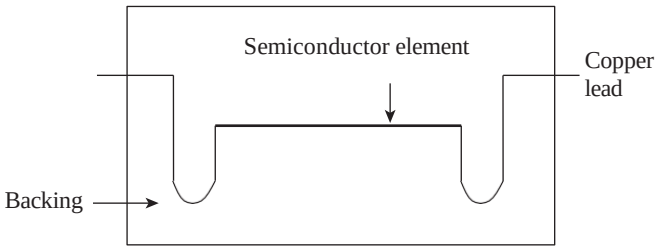


Fig. 14.16 Semiconductor gauge

the electrical resistance of the gauge decreases in response to tensile strain, it is called a negative or n-type semiconductor gauge. A semiconductor gauge is shown in Fig. 14.16.

A semiconductor gauge comprises a rectangular filament having a dimension of 10 × 0.25 × 0.05 mm. It is essential to make them as thin as possible because breaking stresses of the material increase as its cross-sectional area decreases. Without fracturing the gauge, it can be bent to a much smaller radius of curvature. These have certain advantages: they show high strain sensitivity and can measure small strain variations accurately. Semiconductor gauges also possess certain disadvantages:

1. The output obtained in a semiconductor gauge is non-linear in nature.
2. The strain sensitivity of the gauge is temperature dependent.
3. It is more fragile when compared to a wire or foil element.
4. It is more expensive than an ordinary metallic gauge.

14.6 STRAIN GAUGE MATERIALS

Strain gauge sensing materials can be divided into two categories: metallic and semiconductor materials. The following are the five most popular metals used in the manufacture of strain gauges:

Constantan It is an alloy of copper and nickel having a low and controllable temperature coefficient. It finds applications in static and dynamic measurements.

Nichrome It is a nickel–chrome alloy that finds applications in static and dynamic measurements.

Dynaloy It is made up of nickel and iron, which has a high gauge factor, and exhibits high resistance to fatigue. It can be used for dynamic strain measurements.

Stabiloy It is a modified nickel–chrome alloy. It has a wide compensation range, from cryogenic to elevated temperatures, and good fatigue life.

Platinum alloy It is an alloy of platinum and tungsten and displays unusual stability and fatigue life at elevated temperatures. It is used for dynamic and static tests up to a temperature of 800 and 650 °C, respectively.

A specially processed silicon crystal, which is cut into filaments, is used for manufacturing semiconductor gauges. The filaments exhibit either p (positive) or n (negative) characteristics. When compared to metallic sensing elements, the gauge factor is much higher in semiconductor materials. Semiconductor materials are highly sensitive and provide a high output.

Resistances of commercially available strain gauges vary from 60 to 5000 Ω. However, the strain gauge of 120 Ω is considered to be a standard because most of the commercially available strain-measuring equipment are designed for 120 Ω. The safe current carrying range is between 25 and 50 mA. If stress distribution of the specimen is reasonably uniform, then a narrow gauge of 10–

30 mm length can be used. For high strain applications, gauges of shorter length can be employed.

Selection of grid materials is important and is based on a compromise of the following desirable characteristics. The grid materials should have all or many of the following characteristics:

1. High gauge factor
2. High resistivity
3. Low temperature sensitivity
4. High electrical stability
5. High yield point
6. High endurance limit
7. Good workability
8. Good solderability or weldability
9. Low hysteresis
10. Low thermal emf when joined with other materials
11. Good corrosion resistances

14.7 BACKING OR CARRIER MATERIALS

Backing materials provide support, dimensional stability, and some mechanical protection for fragile and delicate strain-sensing elements. Basically, two types of carrier materials are used.

Permanent carriers In addition to supporting the gauge, permanent carriers provide electrical insulation between the grid and the surface of the part under study. These carriers are manufactured from organic materials. They are thin and flexible enough to sustain mechanical contact with the specimen irrespective of the surface contour, temperature gradient, or transient conditions.

Temporary carriers Organic carriers may not be appropriate for certain environmental conditions. In such cases, temporary carriers are employed. Gauges that are used for high-temperature applications employ temporary carriers, which provide temporary backing. The backing is removed once the grid is mounted. The advantage of using temporary carriers is that even though there is an increase in specimen stiffness, there is an improvement in mechanical and thermal contacts. The following are the desirable characteristics of a backing material:

1. The backing material should be very thin.
2. It should have high mechanical strength so that it is strong enough to transmit force from the structure under study to the sensing resistance wire.
3. It should have high dielectric strength, that is, it should be an insulator of electricity.
4. The backing material should not absorb humidity, that is, it should be non-hygroscopic.
5. It should be compatible with the adhesive material used to bond it to the structure under study.
6. It should have minimum temperature restriction, that is, temperature changes should not affect it.

The most commonly used backing materials are thin papers, phenolic-impregnated papers, epoxy-type plastic films, and epoxy-impregnated fibre glass. Most foil gauges intended for a moderate range of temperature (−100 to 200 °F) employ an epoxy film backing.