ces, and their

5 Measurement

Chapter Contents

In addition to mechanical and physical properties of materials, other factors that determine the performance of a manufactured product include the dimensions and surfaces of its components. ***Dimensions*** are the linear or angular sizes of a component specified on the part drawing. Dimensions are important because they determine how well the components of a product fit together during assembly. When fabricating a given component, it is nearly impossible and very costly to make the part to the exact dimension given on the drawing. Instead a limited variation is allowed from the dimension, and that allowable variation is called a ***tolerance***.

The surfaces of a component are also important. They affect product performance, assembly fit, and aesthetic appeal that a potential customer might have for the product. A ***surface*** is the exterior boundary of an object with its surroundings, which may be another object, a fluid, or space, or combinations of these. The surface encloses the object's bulk mechanical and physical properties.

This chapter covers dimensions, tolerances, and surfaces—three attributes specified by the product designer and determined by the manufacturing processes that make the parts and products. It also considers how these attributes are assessed using measuring and gaging devices. Closely related topics are quality control and inspection, covered in Chapter 40.

5. Dimensions, Tolerances, and Related Attributes

The basic parameters used by design engineers to specify sizes of geometric features on a part drawing are defined in this section. The parameters include dimensions and tolerances, flatness, roundness, and angularity.

5.1.1 DIMENSIONS AND TOLERANCES

ANSI [3] defines a ***dimension*** as "a numerical value expressed in appropriate units of measure and indicated on a drawing and in other documents along with lines, symbols, and notes to define the size or geometric characteristic, or both, of a part or part feature." Dimensions on part drawings represent nominal or basic sizes of the part and its features. These are the values that the designer would like the part size to be, if the part could be made to an exact size with no errors or variations in the fabrication process. However, there are variations in the manu- facturing process, which are manifested as variations in the part size. Tolerances are used to define the limits of the allowed variation. Quoting again from the ANSI standard [3], a ***tolerance*** is "the total amount by which a specific dimension is permitted to vary. The tolerance is the difference between the maximum and minimum limits."

Tolerances can be specified in several ways, illustrated in Figure 5.1. Probably most common is the ***bilateral tolerance***, in which the variation is permitted in both positive and negative directions from the nominal dimension. For example, in Figure 5.1(a), the nominal dimension = 2.500 linear units (e.g., mm, in), with an allowable variation of 0.005 units in either direction. Parts outside these limits are unacceptable. It is pos- sible for a bilateral tolerance to be unbalanced; for example, 2.500 + 0.010, −0.005 dimensional units. A ***unilateral tolerance*** is one in which the variation from the speci- fied dimension is permitted in only one direction, either positive or negative, as in Figure 5.1(b). ***Limit dimensions*** are an alternative method to specify the permissible variation in a part feature size; they consist of the maximum and minimum dimensions allowed, as in Figure 5.1(c).

5.1.2 OTHER GEOMETRIC ATTRIBUTES

Dimensions and tolerances are normally expressed as linear (length) values. There are other geometric attributes of parts that are also important, such as flatness of a surface, roundness of a shaft or hole, parallelism between two surfaces, and so on. Definitions of these terms are listed in Table 5.1.

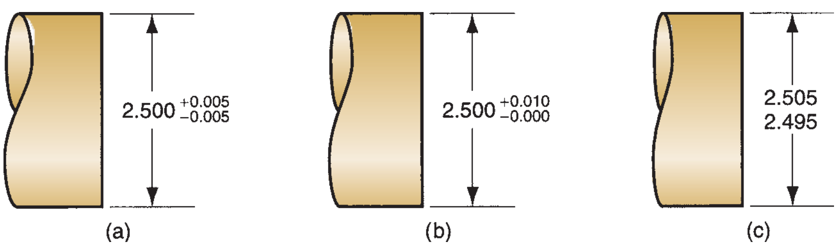


FIGURE 5.1 Three ways to specify tolerance limits for a nominal dimension of 2.500: (a) bilateral, (b) unilateral, and (c) limit dimensions.

TABLE • 5.1 Definitions of geometric attributes of parts.

Angularity—The extent to which a part feature such as a surface or axis is at a specified angle relative to a reference surface. If the angle = 90°, then the attribute is called perpendicularity or squareness.

Circularity—For a surface of revolution such as a cylinder, circular hole, or cone, circularity is the degree to which all points on the intersection of the surface and a plane perpendicular to the axis of revolution are equidistant from the axis. For a sphere, circularity is the degree to which all points on the intersection of the surface and a plane passing through the center are equidistant from the center.

Concentricity—The degree to which any two (or more) part features such as a cylindrical surface and a circular hole have a common axis.

Cylindricity—The degree to which all points on a surface of revolution such as a cylinder are equidistant from the axis of revolution.

Flatness—The extent to which all points on a surface lie in a single plane.

Parallelism—The degree to which all points on a part feature such as a surface, line, or axis are equidistant from a reference plane or line or axis.

Perpendicularity—The degree to which all points on a part feature such as a surface, line, or axis are 90° from a reference plane or line or axis.

Roundness—Same as circularity.

Squareness—Same as perpendicularity.

Straightness—The degree to which a part feature such as a line or axis is a straight line.

5.2 Measuring Instruments and Gages

Measurement is a procedure in which an unknown quantity is compared with a known standard, using an accepted and consistent system of units. Two systems of units have evolved in the world: (1) the U.S. customary system (U.S.C.S.), and (2) the International System of Units (or SI, for Systeme Internationale d'Unites), more popularly known as the metric system (Historical Note 5.1). Both systems are used in parallel throughout this book. The metric system is widely accepted in nearly every part of the industrialized world except the United States, which has stubbornly clung to its U.S.C.S. Gradually, the United States is adopting SI.

Historical Note 5.1 *Measurement Systems*

Measurement systems in ancient civilizations were based on dimensions of the human body. Egyptians developed the cubit as a linear measurement standard around 3000 BCE. The ***cubit*** was defined as the length of a human arm and hand from elbow to fingertip. Although there were difficulties due to variations in arm lengths, the cubit was standardized in the form of a master cubit made of granite. This standard cubit, which measured 524 mm (26.6 in), was used to produce other cubit sticks throughout Egypt. The standard cubit was divided into ***digits*** (a human finger width), with 28 digits to a cubit. Four digits equaled a ***palm***, and five a ***hand***. Thus, a system of measures and standards was developed in the ancient world.

Ultimately, domination of the ancient Mediterranean world passed to the Greeks and then the Romans. The basic linear measure of the Greeks was the ***finger*** (about 19 mm or 3/4 in), and 16 fingers equaled one ***foot.*** The Romans adopted and adapted the Greek system, specifically the foot, dividing it into 12 parts (called ***unciae***). The Romans defined five feet as a ***pace*** and 5000 feet as a ***mile*** (such a nice round number, how did we end up with 5280 feet in a mile?).

During medieval Europe, various national and regional measurement systems developed, many of them based on the Roman standards. Two primary systems emerged in the Western world, the English system and the metric system. The English system defined the ***yard*** "as the distance from the thumb tip to the end of the nose of King Henry I" [15]. The yard was divided into three ***feet*** and one foot into 12 ***inches***. The American colonies were tied to England,

and it was therefore natural for the United States to adopt the same system of measurements at the time of its independence. This became the U.S. customary system (U.S.C.S.).

The initial proposal for the metric system is credited to vicar G. Mouton of Lyon, France around 1670. His proposal included three important features that were subsequently incorporated into the metric standards: (1) the basic unit was defined in terms of a measurement of the Earth, which was presumed constant; (2) the units were subdivided decimally; and (3) units had rational prefixes. Mouton's proposal was discussed and debated among scientists in France for the next 125 years. A result of the French Revolution was the adoption of the metric system of weights and measures (in 1795). The basic unit of length was the ***meter***, which was then defined as 1/10,000,000 of the length of the meridian between the North Pole and the Equator, and passing through Paris (but of course). Multiples and subdivisions of the meter were based on Greek prefixes.

Dissemination of the metric system throughout Europe during the early 1800s was encouraged by the military successes of French armies under Napoleon. In other parts of the world, adoption of the metric system occurred over many years and was often motivated by significant political changes. This was the case in Japan, China, the Soviet Union (Russia), and Latin America.

An act of British Parliament in 1963 redefined the English system of weights and measures in terms of metric units and mandated a changeover to metric two years later, thus aligning Britain with the rest of Europe and leaving the United States as the only major industrial nation that was nonmetric. In 1960, an international conference on weights and measures in Paris reached agreement on new standards based on the metric system. Thus the metric system became the Systeme Internationale (SI).

Measurement provides a numerical value of the quantity of interest, within certain limits of accuracy and precision. ***Accuracy*** is the degree to which the measured value agrees with the true value of the quantity of interest. A measurement procedure is accurate when it is absent of systematic errors, which are positive or negative deviations from the true value that are consistent from one measurement to the next. ***Precision*** is the degree of repeatability in the measurement process. Good precision means that random errors in the measurement procedure are minimized. Random errors are usually associated with human participation in the measurement process. Examples include variations in the setup, imprecise reading of the scale, round-off approximations, and so on. Nonhuman contributors to random error include temperature changes, gradual wear and/or misalignment in the working elements of the device, and other variations.

Closely related to measurement is gaging. ***Gaging*** (also spelled ***gauging***) determines simply whether the part characteristic meets or does not meet the design specification. It is usually faster than measuring, but scant information is provided about the actual value of the characteristic of interest.

This section considers the variety of manually operated measuring instruments and gages used to evaluate dimensions such as length and diameter, as well as features such as angles, straightness, and roundness. This type of equipment is found in metrology labs, inspection departments, and tool rooms. The logical starting topic is precision gage blocks.

5.2.1 PRECISION GAGE BLOCKS

Precision gage blocks are the standards against which other dimensional measuring instruments and gages are compared. Gage blocks are usually square or rectangular. The measuring surfaces are finished to be dimensionally accurate and parallel to within several millionths of an inch and are polished to a mirror finish. Several grades of precision gage blocks are available, with closer tolerances for higher precision grades. The highest grade—the ***master laboratory standard***—is made to a tolerance of ±0.000,03 mm (±0.000,001 in). Depending on degree of hardness desired and price the user is willing to pay, gage blocks can be made out of any of several hard materi- als, including tool steel, chrome-plated steel, chromium carbide, or tungsten carbide.

Precision gage blocks are available in certain standard sizes or in sets, the latter containing a variety of different-sized blocks. The sizes in a set are systematically determined so they can be stacked to achieve virtually any dimension desired to within 0.0025 mm (0.0001 in).

For best results, gage blocks must be used on a flat reference surface, such as a surface plate. A ***surface plate*** is a large solid block whose top surface is finished to a flat plane. Most surface plates today are made of granite. Granite has the advantage of being hard, non-rusting, nonmagnetic, long wearing, thermally stable, and easy to maintain.

Gage blocks and other high-precision measuring instruments must be used under standard conditions of temperature and other factors that might adversely affect the measurement. By international agreement, 20°C (68°F) has been established as the standard temperature. Metrology labs operate at this standard. If gage blocks or other measuring instruments are used in a factory environment in which the temperature differs from this standard, corrections for thermal expansion or contraction may be required. Also, working gage blocks used for inspection in the shop are subject to wear and must be calibrated periodically against more precise laboratory gage blocks.

5.2.2 MEASURING INSTRUMENTS FOR LINEAR DIMENSIONS

Measuring instruments can be divided into two types: graduated and nongraduated. ***Graduated measuring devices*** include a set of markings (called ***graduations***) on a linear or angular scale to which the object's feature of interest can be compared for measurement. ***Nongraduated measuring devices*** possess no such scale and are used to make comparisons between dimensions or to transfer a dimension for measurement by a graduated device.

The most basic of the graduated measuring devices is the ***rule*** (made of steel, and often called a ***steel rule***), used to measure linear dimensions. Rules are available in various lengths. Metric rule lengths include 150, 300, 600, and 1000 mm, with graduations of 1 or 0.5 mm. Common U.S. sizes are 6, 12, and 24 in, with graduations of 1/32, 1/64, or 1/100 in.

Calipers are available in either nongraduated or graduated styles. A nongraduated caliper (referred to simply as a ***caliper***) consists of two legs joined by a hinge mechanism, as in Figure 5.2. The ends of the legs are made to contact the surfaces of the object being measured, and the hinge is designed to hold the legs in position during use. The contacts point either inward or outward. When they point inward, as in Figure 5.2, the instrument is an ***outside caliper*** and is used for measuring outside dimensions such as a diameter. When the contacts point outward, it is an ***inside caliper***, which is used to measure the distance between two internal surfaces. An instrument similar in configuration to the caliper is a ***divider***, except that both legs are straight and terminate in hard, sharply pointed contacts. Dividers are used for scaling distances between two points or lines on a surface, and for scribing circles or arcs onto a surface.

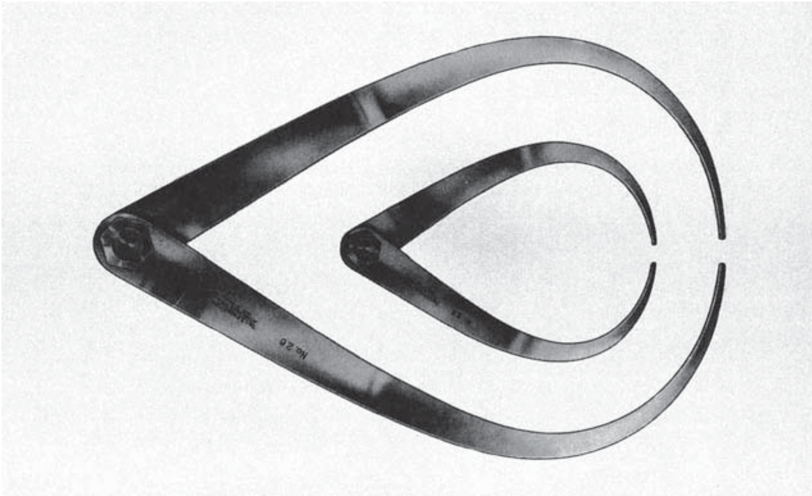


FIGURE 5.2 Two sizes of outside calipers. (Courtesy of L. S. Starrett Co.)

A variety of graduated calipers are available for various measurement purposes. The simplest is the ***slide caliper***, which consists of a steel rule to which two jaws are added, one fixed at the end of the rule and the other movable, shown in Figure 5.3. Slide calipers can be used for inside or outside measurements, depending on whether the inside or outside jaw faces are used. In use, the jaws are forced into contact with the part surfaces to be measured, and the location of the movable jaw indicates the

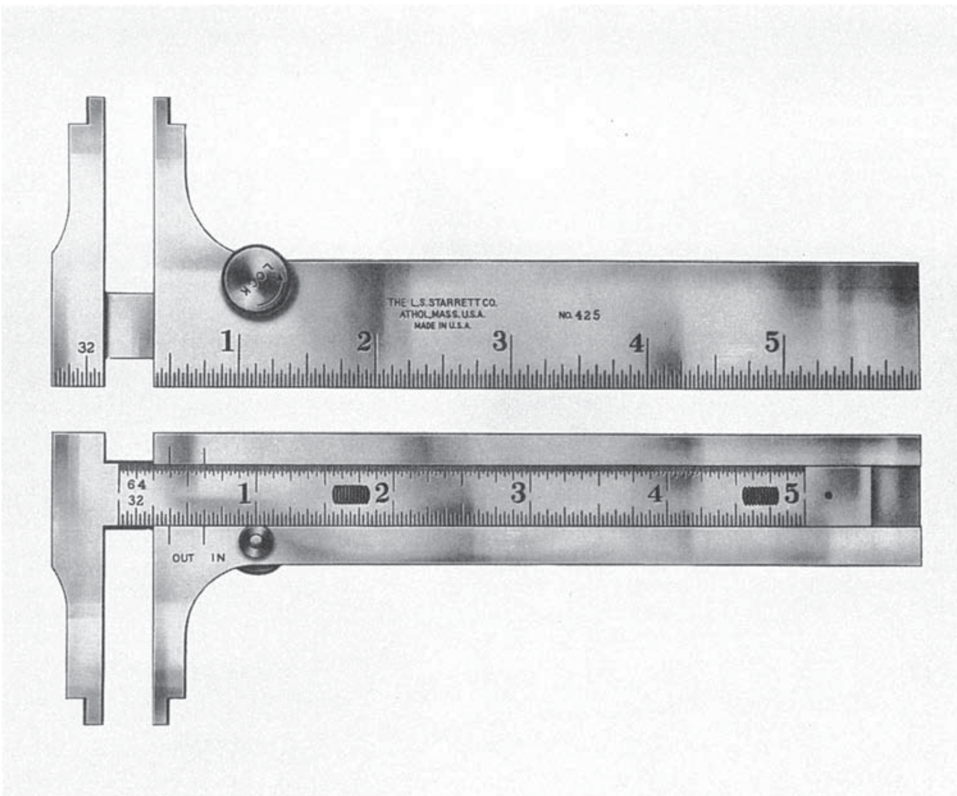


FIGURE 5.3 Slide caliper, opposite sides of instrument shown. (Courtesy of L. S. Starrett Co.)

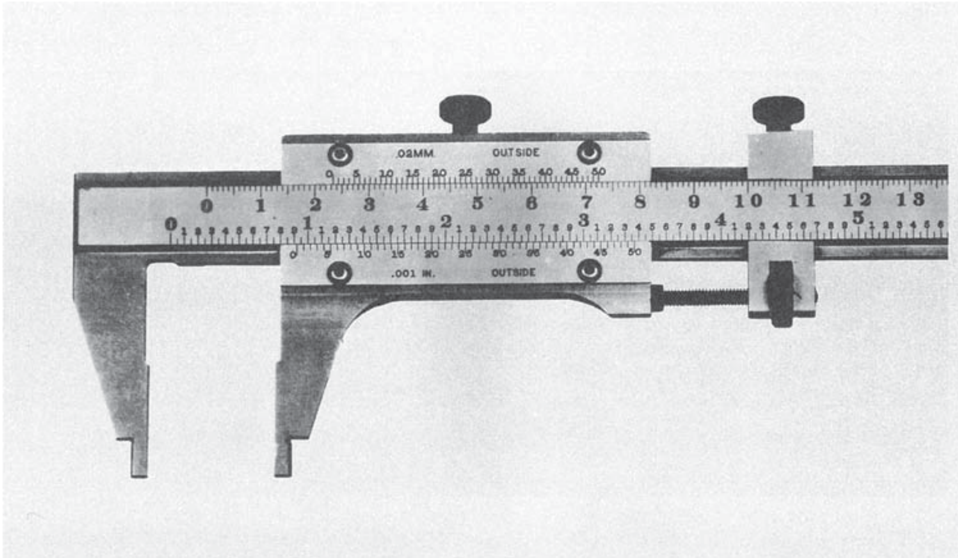


FIGURE 5.4 Vernier caliper. (Courtesy of L. S. Starrett Co.)

dimension of interest. Slide calipers permit more accurate and precise measurements than simple rules. A refinement of the slide caliper is the ***vernier caliper***, shown in Figure 5.4. In this device, the movable jaw includes a vernier scale, named after P. Vernier (1580–1637), a French mathematician who invented it. The vernier provides graduations of 0.01 mm in the SI (and 0.001 inch in the U.S. customary scale), much more precise than the slide caliper.

The ***micrometer*** is a widely used and very accurate measuring device, the most common form of which consists of a spindle and a C-shaped anvil, as in Figure 5.5. The spindle is moved relative to the fixed anvil by means of an accurate screw thread. On a typical U.S. micrometer, each rotation of the spindle provides 0.025 in of linear travel. Attached to the spindle is a thimble graduated with 25 marks around its circumference,

FIGURE 5.5 External micrometer, standard 1-in size with digital readout. (Courtesy of L. S. Starrett Co.)

each mark corresponding to 0.001 in. The micrometer sleeve is usually equipped with a vernier, allowing resolutions as close as 0.0001 in. On a micrometer with metric scale, graduations are 0.01 mm. Modern micrometers (and graduated calipers) are available with electronic devices that display a digital readout of the measurement (as in the figure). These instruments are easier to read and eliminate much of the human error associated with reading conventional graduated devices.

The most common micrometer types are (1) ***external micrometer***, Figure 5.5, also called an ***outside micrometer***, which comes in a variety of standard anvil sizes; (2) ***internal micrometer***, or ***inside micrometer***, which consists of a head assembly and a set of rods of different lengths to measure various inside dimensions that might be encountered; and (3) ***depth micrometer***, similar to an inside micrometer but adapted to measure hole depths.

5.2.3 *COMPARATIVE INSTRUMENTS*

Comparative instruments are used to make dimensional comparisons between two objects, such as a work part and a reference surface. They are usually not capable of providing an absolute measurement of the quantity of interest; instead, they measure the magnitude and direction of the deviation between two objects. Instruments in this category include mechanical and electronic gages.

Mechanical Gages: Dial Indicators ***Mechanical gages*** are designed to mechanically magnify the deviation to permit observation. The most common instrument in this category is the ***dial indicator***, Figure 5.6, which converts and amplifies the linear movement of a contact pointer into rotation of a dial needle. The dial is graduated in

FIGURE 5.6 Dial indicator showing dial and graduated face. (Courtesy of L. S. Starrett Co.)

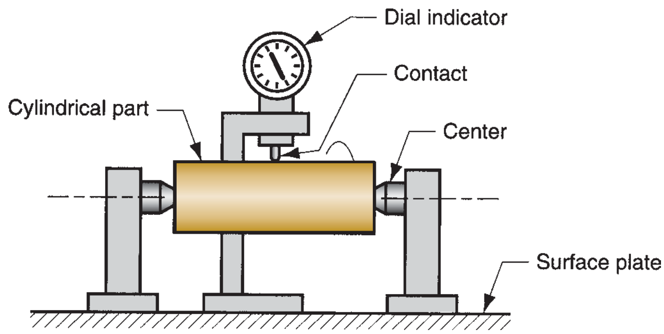


FIGURE 5.7 Dial indicator setup to measure runout; as part is rotated about its center, variations in outside surface relative to center are indicated on the dial.

small units such as 0.01 mm (or 0.001 in). Dial indicators are used in many applications to measure straightness, flatness, parallelism, squareness, roundness, and run-out. A typical setup for measuring run-out is illustrated in Figure 5.7.

Electronic Gages Electronic gages are a family of measuring and gaging instruments based on transducers capable of converting a linear displacement into an electrical signal. The electrical signal is then amplified and transformed into a suitable data format such as a digital readout, as in Figure 5.5. Applications of electronic gages have grown rapidly in recent years, driven by advances in microprocessor technology. They are gradually replacing many of the conventional measuring and gaging devices. Advantages of electronic gages include (1) good sensitivity, accuracy, precision, repeatability, and speed of response; (2) ability to sense very small dimensions—down to 0.025 *m*m (1 *m*-in.); (3) ease of operation; (4) reduced human error; (5) electrical signal that can be displayed in various formats; and (6) capabil- ity to be interfaced with computer systems for data processing.

5.2.4 *FIXED GAGES*

A fixed gage is a physical replica of the part dimension to be assessed. There are two basic categories: master gage and limit gage. A ***master gage*** is fabricated to be a direct replica of the nominal size of the part dimension. It is generally used for setting up a comparative measuring instrument, such as a dial indicator; or for calibrating a measuring device.

A ***limit gage*** is fabricated to be a reverse replica of the part dimension and is designed to check the dimension at one or more of its tolerance limits. A limit gage often consists of two gages in one piece, the first for checking the lower limit of the tolerance on the part dimension, and the other for checking the upper limit. These gages are popularly known as ***GO/NO-GO gages***, because one gage limit allows the part to be inserted, whereas the other limit does not. The ***GO limit*** is used to check the dimension at its maximum material condition; this is the minimum size for an internal feature such as a hole, and it is the maximum size for an external fea- ture such as an outside diameter. The ***NO-GO limit*** is used to inspect the minimum material condition of the dimension in question.

Common limit gages are snap gages and ring gages for checking outside part dimensions, and plug gages for checking inside dimensions. A ***snap gage*** consists of a C-shaped frame with gaging surfaces located in the jaws of the frame, as in

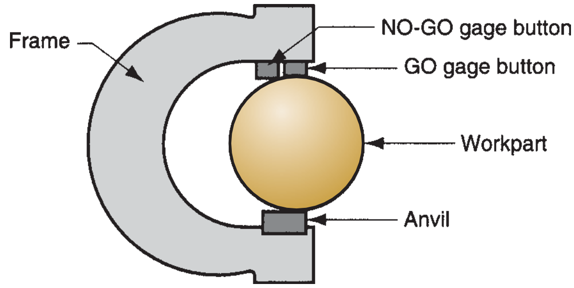


FIGURE 5.8 Snap gage for measuring diameter of a part; difference in height of GO and NO-GO gage buttons is exaggerated.

Figure 5.8. It has two gage buttons, the first being the GO gage, and the second being the NO-GO gage. Snap gages are used for checking outside dimensions such as diameter, width, thickness, and similar surfaces.

Ring gages are used for checking cylindrical diameters. For a given application, a pair of gages is usually required, one GO and the other NO-GO. Each gage is a ring whose opening is machined to one of the tolerance limits of the part diameter. For ease of handling, the outside of the ring is knurled. The two gages are distinguished by the presence of a groove around the outside of the NO-GO ring.

The most common limit gage for checking hole diameter is the ***plug gage***. The typical gage consists of a handle to which are attached two accurately ground cylindrical pieces (plugs) of hardened steel, as in Figure 5.9. The cylindrical plugs serve as the GO and NO-GO gages. Other gages similar to the plug gage include ***taper gages***, consisting of a tapered plug for checking tapered holes; and ***thread gages***, in which the plug is threaded for checking internal threads on parts.

A wear allowance is usually added to the GO limit (maximum material condition) of fixed gages, especially ones that are used frequently. This allowance is expected to account for gradual wearing of the gage GO surface against the part surfaces that are being checked. The wear allowance is typically specified as a percent that is applied to the total tolerance band of the part dimension of interest.

Example 5.1 Wear allowance on a fixed gage

A GO/NO GO plug gage is designed to check a hole diameter that is dimensioned 20.00 mm ± 0.10 mm. A wear allowance of 2.5% of the total tolerance band is applied to the GO side of the gage. Determine the nominal sizes of the GO and NO GO sides of the gage.

Solution: The total tolerance band = 0.10 + 0.10 = 0.20 mm. The wear allowance = 0.025(0.20) = 0.005 mm. The GO gage is used to check the minimum acceptable hole diameter, which is 20.00 − 0.10 = 19.90 mm. Because this is the surface that will wear, the wear allowance is applied to it. Accordingly, the nominal size of the GO gage = 19.90 + 0.005 = 19.905 mm.

The NO GO gage is used to check the maximum hole diameter. Because insertion of the gage will occur only in an out-of-tolerance situation, its wear should be negligible. Therefore, the nominal size of the NO GO gage = 20.00 + 0.10 = 20.10 mm.

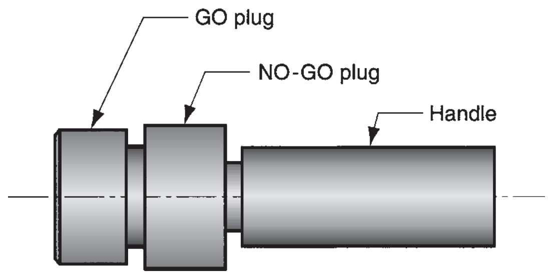


FIGURE 5.9 Plug gage; difference in diameters of GO and NO-GO plugs is exaggerated.

Fixed gages are easy to use, and the time required to complete an inspection is almost always less than when a measuring instrument is employed. Fixed gages were a fundamental element in the development of interchangeable parts manufacturing (Historical Note 1.1). They provided the means by which parts could be made to tolerances that were sufficiently close for assembly without filing and fitting. Their disadvantage is that they provide little if any information on the actual part size; they only indicate whether the size is within tolerance. Today, with the availability of high-speed electronic measuring instruments, and with the need for statistical proc- ess control of part sizes, use of gages is gradually giving way to instruments that provide actual measurements of the dimension of interest.

5.2.5 *ANGULAR MEASUREMENTS*

Angles can be measured using any of several styles of ***protractor***. A ***simple protractor*** consists of a blade that pivots relative to a semicircular head that is graduated in angular units (e.g., degrees, radians). To use, the blade is rotated to a position corresponding to some part angle to be measured, and the angle is read off the angular scale. A ***bevel protractor***, Figure 5.10, consists of two straight blades that

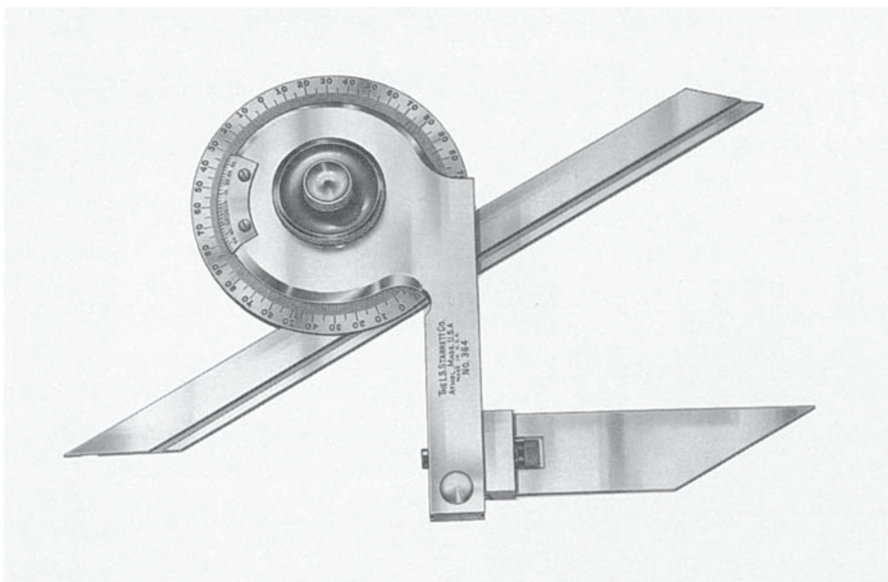


FIGURE 5.10 Bevel protractor with vernier scale. (Courtesy of L. S. Starrett Co.)

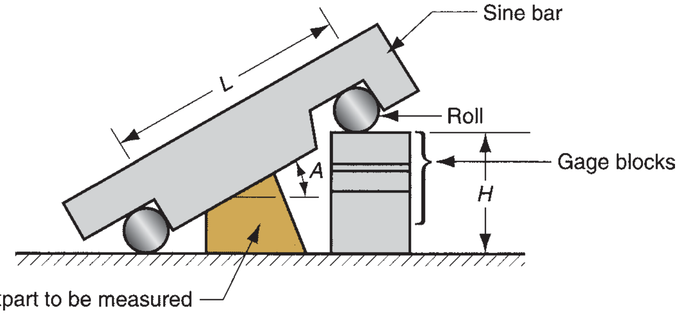


FIGURE 5.11 Setup for using a sine bar.

pivot relative to each other. The pivot assembly has a protractor scale that permits the angle formed by the blades to be read. When equipped with a vernier, the bevel protractor can be read to about 5 min; without a vernier the resolution is only about 1 degree.

High precision in angular measurements can be made using a ***sine bar***, illustrated in Figure 5.11. One possible setup consists of a flat steel straight edge (the sine bar), and two precision rolls set a known distance apart on the bar. The straight edge is aligned with the part angle to be measured, and gage blocks are used or other accurate linear measurements are made to determine height. The proce- dure is carried out on a surface plate to achieve the most accurate results. This height H and the length L of the sine bar between rolls are used to calculate the angle A using

$$\sin A = \frac{H}{L} \tag{5.2}$$

Example 5.2 Sine bar measurement

A sine bar that is 200.00 mm long is used to measure the angle on a part in a setup similar to that in Figure 5.11. Gage blocks are stacked to a height of 40.380 mm. Determine the angle of interest.

Solution: $\sin A = \dfrac{40.38}{200.00} = 0.2019$, $A = \sin^{-1}(0.2019) = 11.64°$.

5.3 Surfaces

A surface is what one touches when holding an object such as a manufactured part. The designer specifies the part dimensions, relating the various surfaces to each other. These ***nominal surfaces***, representing the intended surface contour of the part, are defined by lines in the engineering drawing. The nominal surfaces appear as absolutely straight lines, ideal circles, round holes, and other edges and surfaces that are geometrically perfect. The actual surfaces of a manufactured part are deter- mined by the processes used to make it. The variety of processes available in manu- facturing result in wide variations in surface characteristics, and it is important for engineers to understand the technology of surfaces.

Surfaces are commercially and technologically important for a number of reasons, different reasons for different applications: (1) Aesthetic reasons—surfaces that are smooth and free of scratches and blemishes are more likely to give a favorable impression to the customer. (2) Surfaces affect safety. (3) Friction and wear depend on surface characteristics. (4) Surfaces affect mechanical and physical properties; for example, surface flaws can be points of stress concentration. (5) Assembly of parts is affected by their surfaces; for example, the strength of adhesively bonded joints (Section 30.3) is increased when the surfaces are slightly rough. (6) Smooth surfaces make better electrical contacts.

Surface technology is concerned with (1) defining the characteristics of a surface, (2) surface texture, (3) surface integrity, and (4) the relationship between manufacturing processes and the characteristics of the resulting surface. The first three topics are covered in this section, and the final topic is presented in Section 5.5.

5.3.1 CHARACTERISTICS OF SURFACES

A microscopic view of a part's surface reveals its irregularities and imperfections. The features of a typical surface are illustrated in the highly magnified cross section of the surface of a metal part in Figure 5.12. Although the discussion here is focused on metallic surfaces, these comments apply to ceramics and polymers, with modifications owing to differences in structure of these materials. The bulk of the part, referred to as the ***substrate***, has a grain structure that depends on previous processing of the metal; for example, the metal's substrate structure is affected by its chemical composition, the casting process originally used on the metal, and any deformation operations and heat treatments performed on the casting.

The exterior of the part is a surface whose topography is anything but straight and smooth. In this highly magnified cross section, the surface has roughness, waviness, and flaws. Although not shown here, it also possesses a pattern and/or direction resulting from the mechanical process that produced it. All of these geometric features are included in the term ***surface texture***.

Just below the surface is a layer of metal whose structure differs from that of the substrate. This is called the ***altered layer***, and it is a manifestation of the actions that have been visited upon the surface during its creation and afterward. Manufacturing processes involve energy, usually in large amounts, which operates on the part against its surface. The altered layer may result from work hardening (mechanical energy), heating (thermal energy), chemical treatment, or even electrical energy. The metal in this layer is affected by the application

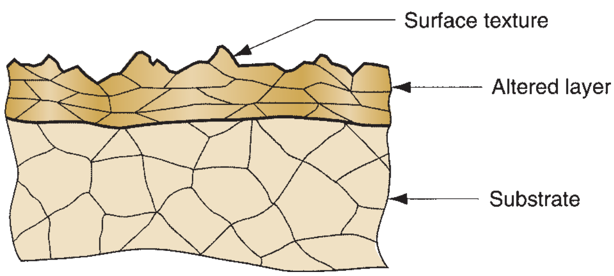


FIGURE 5.12 A magnified cross section of a typical metallic part surface.

of energy, and its microstructure is altered accordingly. This altered layer falls within the scope of ***surface integrity***, which is concerned with the definition, specification, and control of the surface layers of a material (most commonly metals) in manufacturing and subsequent performance in service. The scope of surface integrity is usually interpreted to include surface texture as well as the altered layer beneath.

In addition, most metal surfaces are coated with an ***oxide film***, given sufficient time after processing for the film to form. Aluminum forms a hard, dense, thin film of Al_2O_3 on its surface (which serves to protect the substrate from corrosion), and iron forms oxides of several chemistries on its surface (rust, which provides virtually no protection at all). There is also likely to be moisture, dirt, oil, adsorbed gases, and other contaminants on the part's surface.

5.3.2 SURFACE TEXTURE

Surface texture consists of the repetitive and/or random deviations from the nomi- nal surface of an object; it is defined by four features: roughness, waviness, lay, and flaws, shown in Figure 5.13. ***Roughness*** refers to the small, finely spaced deviations from the nominal surface that are determined by the material characteristics and the process that formed the surface. ***Waviness*** is defined as the deviations of much larger spacing; they occur because of work deflection, vibration, heat treatment, and similar factors. Roughness is superimposed on waviness. ***Lay*** is the predominant direction or pattern of the surface texture. It is determined by the manufacturing method used to create the surface, usually from the action of a cutting tool. Figure 5.14 presents most of the possible lays a surface can take, together with the symbol used by a designer to specify them. Finally, ***flaws*** are irregularities that occur occasionally on the sur- face; these include cracks, scratches, inclusions, and similar defects in the surface. Although some of the flaws relate to surface texture, they also affect surface integrity (Section 5.3.3).

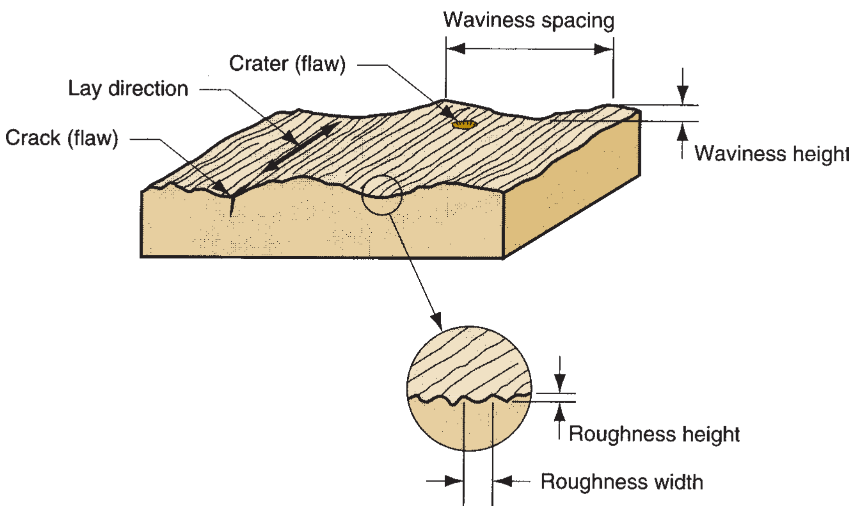


FIGURE 5.13 Surface texture features.

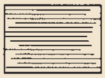
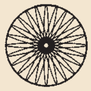
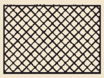
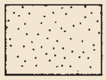
Lay symbol	Surface pattern	Description	Lay symbol	Surface pattern	Description
=		Lay is parallel to line representing surface to which symbol is applied.	C		Lay is circular relative to center of surface to which symbol is applied.
⊥		Lay is perpendicular to line representing surface to which symbol is applied.	R		Lay is approximately radial relative to the center of the surface to which symbol is applied.
X		Lay is angular in both directions to line representing surface to which symbol is applied.	P		Lay is particulate, nondirectional, or protuberant.

FIGURE 5.14 Possible lays of a surface. ([1]).

Surface Roughness and Surface Finish Surface roughness is a measurable characteristic based on the roughness deviations as defined above. ***Surface finish*** is a more subjective term denoting smoothness and general quality of a surface. In popular usage, surface finish is often used as a synonym for surface roughness.

The most commonly used measure of surface texture is surface roughness. With respect to Figure 5.15, ***surface roughness*** can be defined as the average of the vertical deviations from the nominal surface over a specified surface length. An arithmetic average (AA) is generally used, based on the absolute values of the deviations, and this roughness value is referred to by the name ***average roughness***. In equation form,

$$R_a = \int_0^{L_m} \frac{|y|}{L_m}\, dx \tag{5.1}$$

where R_a = arithmetic mean value of roughness, m (in); y = the vertical deviation from nominal surface (converted to absolute value), m (in); and L_m = the specified distance over which the surface deviations are measured. An approximation of Equation (5.1), perhaps easier to comprehend, is given by

$$R_a = \sum_{i=1}^{n} \frac{|y_i|}{n} \tag{5.2}$$

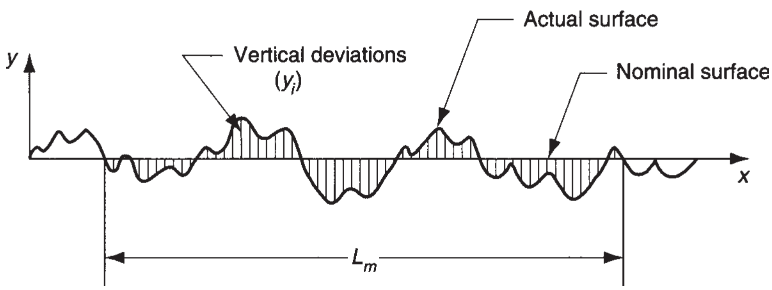


FIGURE 5.15 Deviations from nominal surface used in the two definitions of surface roughness.

where R_a has the same meaning as above; y_i = vertical deviations converted to absolute value and identified by the subscript i, m (in); and n = the number of deviations included in L_m. The units in these equations are meters and inches. In fact, the scale of the deviations is very small, so more appropriate units are mm (mm = m X 10^{-6} = mm X 10^{-3}) or m-in (m-in = inch X 10^{-6}). These are the units commonly used to express surface roughness.

The AA method is the most widely used averaging method for surface roughness today. An alternative, sometimes used in the United States, is the ***root-mean-square*** (RMS) average, which is the square root of the mean of the squared deviations over the measuring length. RMS surface roughness values will almost always be greater than the AA values because the larger deviations will figure more prominently in the calculation of the RMS value.

Surface roughness suffers the same kinds of deficiencies of any single measure used to assess a complex physical attribute. For example, it fails to account for the lay of the surface pattern; thus, surface roughness may vary significantly, depending on the direction in which it is measured.

Another deficiency is that waviness can be included in the R_a computation. To deal with this problem, a parameter called the ***cutoff length*** is used as a filter that separates the waviness in a measured surface from the roughness deviations. In effect, the cutoff length is a sampling distance along the surface. A sampling distance shorter than the waviness width will eliminate the vertical deviations associated with waviness and only include those associated with roughness. The most common cutoff length used in practice is 0.8 mm (0.030 in). The measuring length L_m is normally set at about five times the cutoff length.

The limitations of surface roughness have motivated the development of additional measures that more completely describe the topography of a given surface. These measures include three-dimensional graphical renderings of the surface, as described in [17].

Symbols for Surface Texture Designers specify surface texture on an engineering drawing by means of symbols as in Figure 5.16. The symbol designating surface texture parameters is a check mark (looks like a square root sign), with entries as indicated for average roughness, waviness, cutoff, lay, and maximum roughness spacing. The symbols for lay are from Figure 5.14.

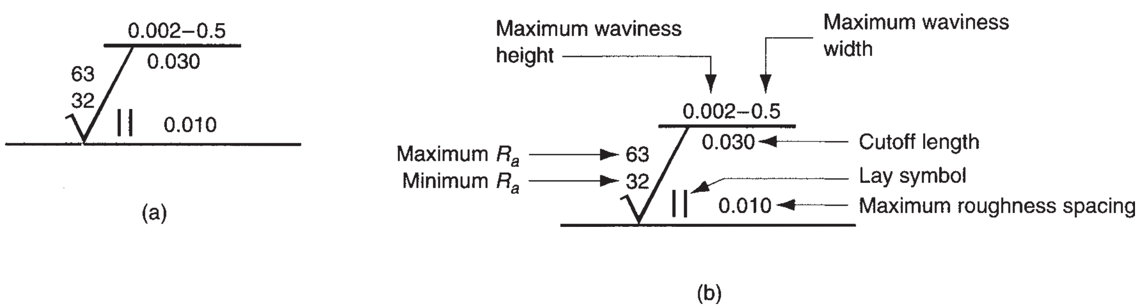


FIGURE 5.16 Surface texture symbols in engineering drawings: (a) the symbol, and (b) symbol with identification labels. Values of R_a are given in microinches; units for other measures are given in inches. Designers do not always specify all of the parameters on engineering drawings.

5.3.3 SURFACE INTEGRITY

Surface texture alone does not completely describe a surface. There may be metallurgical or other changes in the material immediately beneath the surface that can have a significant effect on its mechanical properties. ***Surface integrity*** is the study and control of this subsurface layer and any changes in it as a result of processing that may influence the performance of the finished part or product. This subsurface layer was previously referred to as the altered layer when its structure differs from the substrate, as in Figure 5.12.

The possible alterations and injuries to the subsurface layer that can occur in manufacturing are listed in Table 5.2. The surface changes are caused by the application of various forms of energy during processing—mechanical, thermal, chemical, and electrical. Mechanical energy is the most common form used in manufacturing; it is applied against the work material in operations such as metal forming (e.g., forging, extrusion), pressworking, and machining. Although its primary function in these processes is to change the geometry of the work part, mechanical energy can also cause residual stresses, work hardening, and cracks in the surface layers. Table 5.3 indicates

TABLE • 5.2 Surface and subsurface alterations that define surface integrity.[a]

Absorption are impurities that are absorbed and retained in surface layers of the base material, possibly leading to embrittlement or other property changes.

Alloy depletion occurs when critical alloying elements are lost from the surface layers, with possible loss of properties in the metal.

Cracks are narrow ruptures or separations either at or below the surface that alter the continuity of the material. Cracks are characterized by sharp edges and length-to-width ratios of 4:1 or more. They are classified as macroscopic (can be observed with magnification of 10X or less) and microscopic (requires magnification of more than 10X).

Craters are rough surface depressions left in the surface by short circuit discharges; associated with electrical processing methods such as electric discharge machining and electrochemical machining (Chapter 25).

Hardness changes refer to hardness differences at or near the surface.

Heat affected zone are regions of the metal that are affected by the application of thermal energy; the regions are not melted but are sufficiently heated that they undergo metallurgical changes that affect properties. Abbreviated HAZ, the effect is most prominent in fusion welding operations (Chapter 29).

Inclusions are small particles of material incorporated into the surface layers during processing; they are a discontinuity in the base material. Their composition usually differs from the base material.

Intergranular attack refers to various forms of chemical reaction at the surface, including intergranular corrosion and oxidation.

Laps, ***folds***, ***seams*** are irregularities and defects in the surface caused by plastic working of overlapping surfaces.

Pits are shallow depressions with rounded edges formed by any of several mechanisms, including selective etching or corrosion; removal of surface inclusions; mechanically formed dents; or electrochemical action.

Plastic deformation refers to microstructural changes from deforming the metal at the surface; it results in strain hardening.

Recrystallization involves the formation of new grains in strain hardened metals; associated with heating of metal parts that have been deformed.

Redeposited metal is metal that is removed from the surface in the molten state and then reattached prior to solidification.

Resolidified metal is a portion of the surface that is melted during processing and then solidified without detaching from the surface. The name ***remelted metal*** is also used for resolidified metal. ***Recast metal*** is a term that includes both redeposited and resolidified metal.

Residual stresses are stresses remaining in the material after processing.

Selective etch is a form of chemical attack that concentrates on certain components in the base material.

[a] Compiled from [2].

TABLE • 5.3 Forms of energy applied in manufacturing and the resulting possible surface and subsurface alterations that can occur.[a]

Mechanical	Thermal	Chemical	Electrical
Residual stresses in subsurface layer Cracks—microscopic and macroscopic Plastic deformation Laps, folds, or seams Voids or inclusions Hardness variations (e.g., work hardening)	Metallurgical changes (recrystallization, grain size changes, phase changes at surface) Redeposited or resolidified material Heat-affected zone Hardness changes	Intergranular attack Chemical contamination Absorption of elements such as H and Cl Corrosion, pitting, and etching Dissolving of microconstituents Alloy depletion	Changes in conductivity and/or magnetism Craters resulting from short circuits during certain electrical processing techniques

[a] Based on [2].

the various types of surface and subsurface alterations that are attributable to the different forms of energy applied in manufacturing. Most of the alterations in the table refer to metals, for which surface integrity has been most intensively studied.

5.4 Measurement of Surfaces

Surfaces are described as consisting of two parameters: (1) surface texture and (2) surface integrity. This section is concerned with the measurement of these two parameters.

5.4.1 *MEASUREMENT OF SURFACE ROUGHNESS*

Various methods are used to assess surface roughness. They can be divided into three categories: (1) subjective comparison with standard test surfaces, (2) stylus electronic instruments, and (3) optical techniques.

Standard Test Surfaces Sets of standard surface finish blocks are available, produced to specified roughness values. To estimate the roughness of a given test specimen, the surface is compared with the standard both visually and by the "fingernail test." In this test, the user gently scratches the surfaces of the specimen and the stand- ards, judging which standard is closest to the specimen. Standard test surfaces are a convenient way for a machine operator to obtain an estimate of surface roughness. They are also useful for design engineers in judging what value of surface roughness to specify on a part drawing.

Stylus Instruments The disadvantage of the fingernail test is its subjectivity. Several stylus-type instruments are commercially available to measure surface roughness—similar to the fingernail test, but more scientific. An example is the instrument shown in Figure 5.17 tracing a standard test surface. In these elec- tronic devices, a cone-shaped diamond stylus with point radius of about 0.005 mm (0.0002 in) and 90° tip angle is traversed across the test surface at a constant slow speed. The operation is depicted in Figure 5.18. As the stylus head is traversed horizontally, it also moves vertically to follow the surface deviations. The vertical movement is converted into an electronic signal that represents the topography

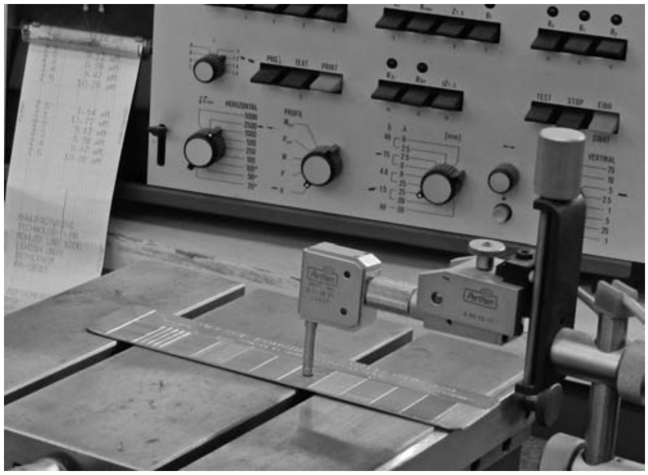


FIGURE 5.17 Stylus-type instrument for measuring surface roughness traversing a standard test surface. (Courtesy of George E. Kane Manufacturing Technology Laboratory, Lehigh University.)

of the surface. This can be displayed as either a profile of the actual surface or an average roughness value. ***Profiling devices*** use a separate flat plane as the nominal reference against which deviations are measured. The output is a plot of the surface contour along the line traversed by the stylus. This type of system can identify both roughness and waviness in the test surface. ***Averaging devices*** reduce the roughness deviations to a single value R_a. They use skids riding on the actual surface to establish the nominal reference plane. The skids act as a mechanical filter to reduce the effect of waviness in the surface; in effect, these averaging devices electronically perform the computations in Equation (5.1).

Optical Techniques Most other surface-measuring instruments employ optical techniques to assess roughness. These techniques are based on light reflectance from the surface, light scatter or diffusion, and laser technology. They are useful in applications where stylus contact with the surface is undesirable. Some of the techniques permit very-high-speed operation, thus making 100% inspection feasible. However, the optical techniques yield values that do not always correlate well with roughness measurements made by stylus-type instruments.

5.4.2 EVALUATION OF SURFACE INTEGRITY

Surface integrity is more difficult to assess than surface roughness. Some of the techniques to inspect for subsurface changes are destructive to the material specimen. Evaluation techniques for surface integrity include the following:

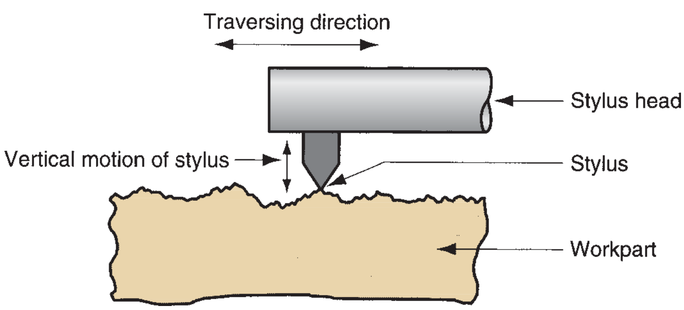


FIGURE 5.18 Sketch illustrating the operation of stylus-type instrument. Stylus head traverses horizontally across surface, while stylus moves vertically to follow surface profile. Vertical movement is converted into either (1) a profile of the surface or (2) the average roughness value.

- ***Surface texture***. Surface roughness, designation of lay, and other measures pro- vide superficial data on surface integrity. This type of testing is relatively simple to perform and is always included in the evaluation of surface integrity.
- ***Visual examination***. Visual examination can reveal various surface flaws such as cracks, craters, laps, and seams. This type of assessment is often augmented by fluorescent and photographic techniques.
- ***Microstructural examination***. This involves standard metallographic techniques for preparing cross sections and obtaining photomicrographs for examination of microstructure in the surface layers compared with the substrate.
- ***Microhardness profile***. Hardness differences near the surface can be detected using microhardness measurement techniques such as Knoop and Vickers (Section 3.2.1). The part is sectioned, and hardness is plotted against distance below the surface to obtain a hardness profile of the cross section.
- ***Residual stress profile***. X-ray diffraction techniques can be employed to measure residual stresses in the surface layers of a part.

5.5 Effect of Manufacturing Processes

The ability to achieve a certain tolerance or surface is a function of the manufacturing process. This section describes the general capabilities of various processes in terms of tolerance and surface roughness and surface integrity.

Some manufacturing processes are inherently more accurate than others. Most machining processes are quite accurate, capable of tolerances of ±0.05 mm (±0.002 in) or better. By contrast, sand castings are generally inaccurate, and tolerances of 10 to 20 times those used for machined parts should be specified. In Table 5.4, we list a variety

TABLE • 5.4 Typical tolerance limits, based on process capability (Section 40.2), for various manufacturing processes.[b]

Process	Typical Tolerance, mm (in)	Process	Typical Tolerance, mm (in)
Sand casting		Abrasive	
Cast iron	±1.3 (±0.050)	Grinding	±0.008 (±0.0003)
Steel	±1.5 (±0.060)	Lapping	±0.005 (±0.0002)
Aluminum	±0.5 (±0.020)	Honing	±0.005 (±0.0002)
Die casting	±0.12 (±0.005)	Nontraditional and thermal	
Plastic molding:		Chemical machining	±0.08 (±0.003)
Polyethylene	±0.3 (±0.010)	Electric discharge	±0.025 (±0.001)
Polystyrene	±0.15 (±0.006)	Electrochem. grind	±0.025 (±0.001)
Machining:		Electrochem. machine	±0.05 (±0.002)
Drilling, 6 mm (0.25 in)	+0.08/-0.03 (+0.003/-0.001)	Electron beam cutting	±0.08 (±0.003)
Milling	±0.08 (±0.003)	Laser beam cutting	±0.08 (±0.003)
Turning	±0.05 (±0.002)	Plasma arc cutting	±1.3 (±0.050)

[b] Compiled from [4], [5], and other sources. For each process category, tolerances vary depending on process parameters. Also, tolerances increase with part size.

TABLE • 5.5 Surface roughness values produced by the various manufacturing processes.[a]

Process	Typical Finish	Roughness Range [b]	Process	Typical Finish	Roughness Range [b]
Casting:			Abrasive:		
Die casting	Good	1–2 (30–65)	Grinding	Very good	0.1–2 (5–75)
Investment	Good	1.5–3 (50–100)	Honing	Very good	0.1–1 (4–30)
Sand casting	Poor	12–25 (500–1000)	Lapping	Excellent	0.05–0.5 (2–15)
Metal forming:			Polishing	Excellent	0.1–0.5 (5–15)
Cold rolling	Good	1–3 (25–125)	Superfinish	Excellent	0.02–0.3 (1–10)
Sheet metal draw	Good	1–3 (25–125)	Nontraditional:		
Cold extrusion	Good	1–4 (30–150)	Chemical milling	Medium	1.5–5 (50–200)
Hot rolling	Poor	12–25 (500–1000)	Electrochemical	Good	0.2–2 (10–100)
Machining:			Electric discharge	Medium	1.5–15 (50–500)
Boring	Good	0.5–6 (15–250)	Electron beam	Medium	1.5–15 (50–500)
Drilling	Medium	1.5–6 (60–250)	Laser beam	Medium	1.5–15 (50–500)
Milling	Good	1–6 (30–250)	Thermal:		
Reaming	Good	1–3 (30–125)	Arc welding	Poor	5–25 (250–1000)
Shaping and planing	Medium	1.5–12 (60–500)	Flame cutting	Poor	12–25 (500–1000)
Sawing	Poor	3–25 (100–1000)	Plasma arc cutting	Poor	12–25 (500–1000)
Turning	Good	0.5–6 (15–250)			

[a] Compiled from [1], [2], and other sources.
[b] Roughness range values are given, *m*m (*m*-in). Roughness can vary significantly for a given process, depending on process parameters.

of manufacturing processes and indicate the typical tolerances for each process. Toler- ances are based on the process capability for the particular manufacturing operation, as defined in Section 40.2. The tolerance that should be specified is a function of part size; larger parts require more generous tolerances. The table lists tolerances for moderately sized parts in each processing category.

The manufacturing process determines surface finish and surface integrity. Some processes are capable of producing better surfaces than others. In general, process- ing cost increases with improvement in surface finish. This is because additional operations and more time are usually required to obtain increasingly better surfaces. Processes noted for providing superior finishes include honing, lapping, polishing, and superfinishing (Chapter 24). Table 5.5 indicates the usual surface roughness that can be expected from various manufacturing processes.