INSPECTION OF METALS

UNDERSTANDING THE BASICS

INSPECTION OF METALS

UNDERSTANDING THE BASICS

Edited by
F.C. Campbell

ASM International®
Materials Park, Ohio 44073-0002
www.asminternational.org

First printing, April 2013

Comments, criticisms, and suggestions are invited, and should be forwarded to ASM International.

Prepared under the direction of the ASM International Technical Book Committee (2012–2013), Bradley J. Diak, *Chair.*

ASM International staff who worked on this project include Scott Henry, Senior Manager, Content Development and Publishing; Karen Marken, Senior Managing Editor; Victoria Burt, Content Developer; Steve Lampman, Content Developer; Sue Sellers, Editorial Assistant; Bonnie Sanders, Manager of Production; Madrid Tramble, Senior Production Coordinator; and Diane Whitelaw, Production Coordinator.

Library of Congress Control Number: 2012955193
ISBN-13: 978-1-62708-000-2
ISBN-10: 0-62708-000-7
SAN: 204-7586

ASM International®
Materials Park, OH 44073-0002
www.asminternational.org

Printed in the United States of America

Contents

Preface

Inspection of metals is used to ensure that the quality of the part or product meets minimum quality and safety requirements. There are hundreds of methods used to inspect metals during its fabrication (in-process inspection), when the part is completed and ready for delivery (final inspection), and during its service life (in-service inspection). While the three stages of inspection are addressed to some extent in this book, the emphasis is on final part inspection. Because it is not possible to address all the different inspection methods used in the industry, only the most widely used inspection methods are covered.

The first half of this book attempts to answer three questions for each of these inspection methods:

- How is the inspection method performed?
- When is it used?
- How does it compare with other inspection methods?

The inspection methods covered are:

- Visual inspection
- Coordinate measuring machines
- Machine vision
- Hardness testing
- Tensile testing
- Chemical composition
- Metallography
- Liquid penetrant, magnetic particle, and eddy current inspection
- Radiographic inspection
- Ultrasonic inspection

The second half of the book covers how these inspection methods are used in different metal fabrication industries:

- Castings
- Steel bar and wire
- Tubular products
- Forgings
- Powder metallurgy parts
- Weldments and brazed assemblies

The emphasis in the second half of the book shows why certain inspection methods are selected for different product forms.

Since the purpose of this book is to cover the basics of inspection of metals, the reader is referred to more advanced texts for detailed information. In particular, for nondestructive test methods, *Nondestructive Evaluation and Quality Control*, Volume 17, *ASM Handbook*, for mechanical property test methods, *Mechanical Testing and Evaluation*, Volume 8, *ASM Handbook*, and for metallography, *Metallography and Microstructures*, Volume 9, *ASM Handbook*.

I would like to acknowledge the help and guidance of Karen Marken, ASM International, and the staff at ASM for their valuable contributions.

F.C. Campbell

CHAPTER 1

Inspection Methods—Overview and Comparison

INSPECTION is an organized examination or formal evaluation exercise. In engineering, inspection involves the measurements, tests, and gages applied to certain characteristics in regard to an object or activity. The results are usually compared to specified requirements and standards for determining whether the item or activity is in line with these targets. Some inspection methods are destructive; however, inspections are usually nondestructive.

Nondestructive examination (NDE), or nondestructive testing (NDT), are a number of technologies used to analyze materials for either inherent flaws (such as fractures or cracks), or damage from use. Some common methods are visual, microscopy, liquid or dye penetrant inspection, magnetic particle inspection, eddy current testing, x-ray or radiographic testing, and ultrasonic testing. This chapter provides an overview of the inspection methods that will be covered in the remainder of this book.

Visual Inspection

Visual inspection provides a means of detecting and examining a variety of surface flaws, such as corrosion, contamination, surface finish, and surface discontinuities on joints (for example, welds, seals, and solder connections). Visual inspection is also the most widely used method for detecting and examining surface cracks that are particularly important because of their relationship to structural failure mechanisms. Even when other inspection techniques are used to detect surface cracks, visual inspection often provides a useful supplement. For example, when the eddy

current examination of process tubing is performed, visual inspection is often performed to verify and more closely examine the surface disturbance. In some instances, acid etching (macroetching) can be used to reveal structures that would not be visible to the naked eye, as shown in the flow lines in Fig. 1.

Given the wide variety of surface flaws that may be detectable by visual examination, the use of visual inspection can encompass different techniques, depending on the product and the type of surface flaw being monitored. The methods of visual inspection involve a wide variety of equipment, ranging from examination with the naked eye to the use of interference microscopes for measuring the depth of scratches in the finish of finely polished or lapped surfaces.

Coordinate Measuring Machines

Coordinate measuring machines (CMMs) are used to inspect the dimensions of a finished product. CMMs consist of the machine itself and its probes and moving arms for providing measurement input, a computer for making rapid calculations and comparisons based on the measurement input, and the computer software that controls the entire system. An example of a CMM probe taking measurements on a machined stiffener is illustrated in Fig. 2. Coordinate measuring machines are primarily charac-

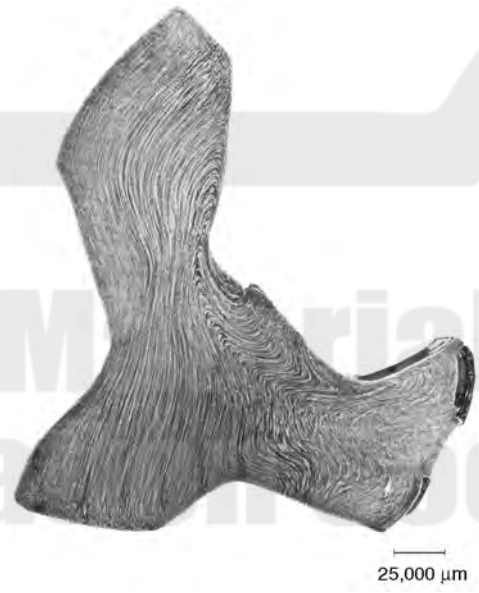


Fig. 1 Flow lines in closed die forged UNS G41400 steering knuckle revealed by cold deep acid etching with 10% aqueous HNO_3 (0.5×) and enhanced with inking. Source: Ref 1

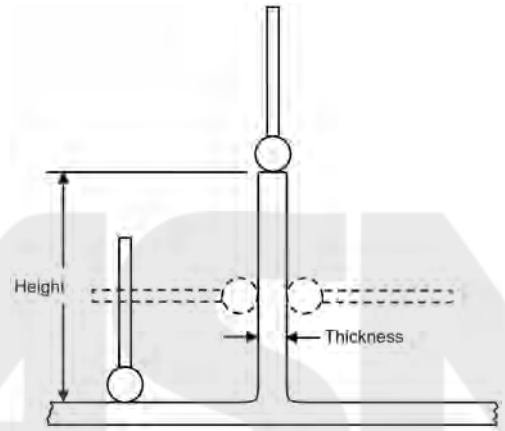


Fig. 2 Example of height and thickness measurements with coordinate measuring machine (CMM) probe

terized by their flexibility, being able to make many measurements without adding or changing tools.

Historically, traditional measuring devices and CMMs have been largely used to collect inspection data on which to make the decision to accept or reject parts. Although CMMs continue to play this role, manufacturers are placing new emphasis on using CMMs to capture data from many sources and bringing them together centrally where they can be used to control the manufacturing process more effectively and preventing defective components from being produced. In addition, CMMs are also being used in entirely new applications; for example, reverse engineering and computer-aided design and manufacture (CAD/CAM) applications as well as innovative approaches to manufacturing, such as the flexible manufacturing systems, manufacturing cells, machining centers, and flexible transfer lines.

Machine Vision

Machine vision emerged as an important new technique for industrial inspection and quality control in the early 1980s. When properly applied, machine vision can provide accurate and inexpensive inspection of workpieces, thus dramatically increasing product quality. Machine vision is also used as an in-process gaging tool for controlling the process and correcting trends that could lead to the production of defective parts. The automotive and electronics industries make heavy use of machine vision for automated high volume, labor intensive and repetitive inspection operations.

This ability to acquire an image, analyze it, and then make an appropriate decision is extremely useful in inspection and quality control applications. It enables machine vision to be used for a variety of functions, including: identification of shapes, measurement of distances and ranges, gaging of sizes and dimensions, determining orientation of parts, quantifying motion, and detecting surface shading. Several examples of machine vision applications are shown in Fig. 3. These capabilities allow users to employ machine vision systems for cost-effective and reliable 100% inspection of workpieces.

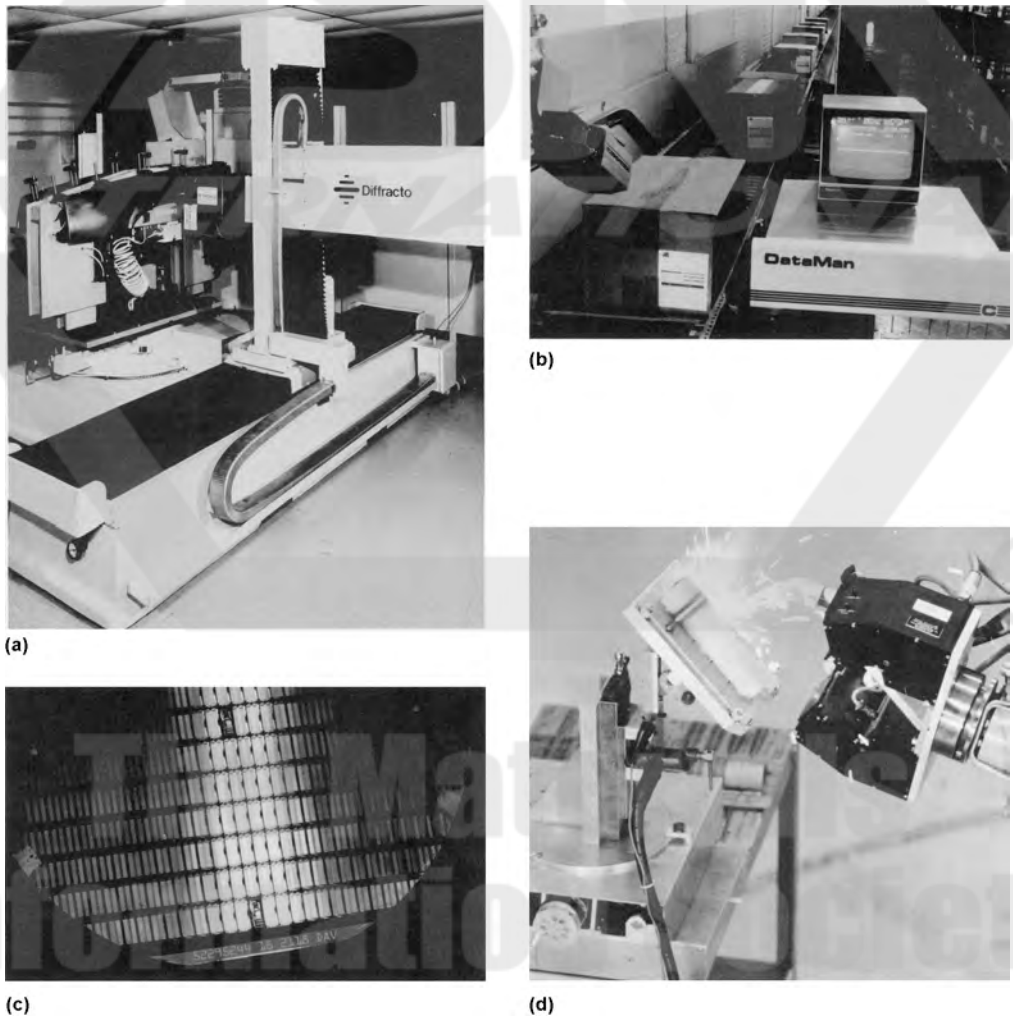


Fig. 3 Examples of machine vision applications. (a) Measuring fit and gap of automotive fender. Courtesy of Diffracto Limited. (b) Reading box labels in sorting application. Courtesy of Cognex Corporation. (c) Reading part numbers on silicon wafers. Courtesy of Cognex Corporation. (d) Vision system for arc welding. Courtesy of Robotic Vision Systems, Inc. Source: Ref 2

Hardness Testing

Hardness testing is one of the simplest and most widely used inspection methods. It is a nondestructive method that can be used to predict the strength of metals. The correlation between tensile strength and hardness for steels, brass, and nodular cast iron are shown in Fig. 4. All heat treated steels are subjected to hardness testing to verify that the heat treatment produced the correct hardness and thus strength.

The most common types of hardness tests are indentation methods. These tests use a variety of indentation loads ranging from 1 gf (microindentation) to 3000 kgf (Brinell). Low and high powered microscopes (Brinell, Vickers, and microindentation) are used to measure the resulting indentation diagonals from which a hardness number is calculated using a formula. In the Rockwell test, the depth of indentation is measured and converted to a hardness number, which is inversely related to the depth.

A general comparison of indentation hardness testing methods is given in Table 1. Generally, the scale to use for a specified material is indicated on the engineering design drawings or in the test specifications. However, at times the scale must be determined and selected to suit a given set of circumstances.

Hardness testing has many applications in quality control, materials evaluation, and the prediction of properties. Because hardness testing is nondestructive and quick, it is a very useful tool for manufacturing and process control. For example, the most common application of the Rockwell test is testing steels that have been hardened and tempered. If a hard-

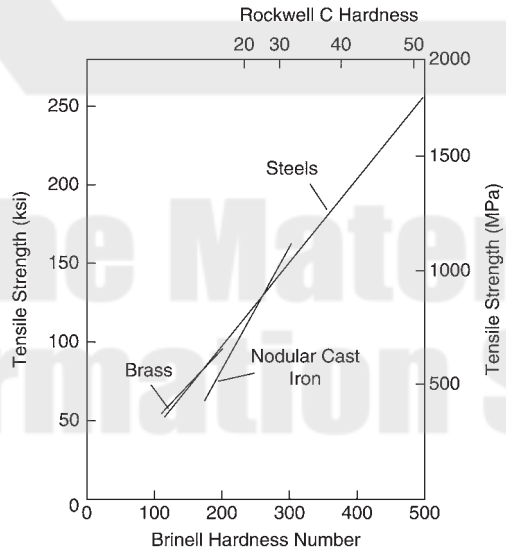


Fig. 4 Correlation of hardness with tensile strength. Source: Ref 3

Table 1 Comparison of indentation hardness tests

The minimum material thickness for a test is usually taken to be 10 times the indentation depth.

		Indent							
Test	Indenter(s)	Diagonal or diameter	Depth	Load(s)	Method of measurement	Surface preparation	Tests per hour	Applications	Remarks
Brinell	Ball indenter, 10 mm (0.4 in.) or 2.5 mm (0.1 in.) in diameter	1–7 mm (0.04–0.28 in.)	Up to 0.3 mm (0.01 in.) and 1 mm (0.04 in.), respectively, with 2.5 mm (0.1 in.) and 10 mm (0.4 in.) diam balls	3000 kgf for ferrous materials down to 100 kgf for soft metals	Measure diameter of indentation under microscope; read hardness from tables	Specially ground area of measurements of diameter	50 with diameter measurements	Large forged and cast parts	Damage to specimen minimized by use of lightly loaded ball indenter. Indent then less than Rockwell
Rockwell	120° diamond cone, 1.6–13 mm ($\frac{1}{16}$ to ½ in.) diam ball	0.1–1.5 mm (0.004–0.06 in.)	25–375 μm (0.1–1.48 μin.)	Major 60–150 kgf Minor 10 kgf	Read hardness directly from meter or digital display	No preparation necessary on many surfaces	300 manually 900 automatically	Forgings, castings, roughly machined parts	Measure depth of penetration, not diameter
Rockwell superficial	As for Rockwell	0.1–0.7 mm (0.004–0.03 in.)	10–110 μm (0.04–0.43 μin.)	Major 15–45 kgf Minor 3 kgf	As for Rockwell	Machined surface, ground	As for Rockwell	Critical surfaces of finished parts	A surface test of case hardening and annealing
Vickers	136° diamond pyramid	Measure diagonal, not diameter	300–100 μm (0.12–0.4 μin.)	1–120 kgf	Measure indent with low-power microscope; read hardness from tables	Smooth clean surface, symmetrical if not flat	Up to 180	Fine finished surfaces, thin specimens	Small indent but high local stresses
Microhardness	136° diamond indenter or a Knoop indenter	40 μm (0.16 μin.)	1–4 μm (0.004–0.016 μin.)	1 gf–1 kgf	Measure indentation with low-power microscope; read hardness from tables	Polished surface	Up to 60	Surface layers, thin stock, down to 200 μm	Laboratory test used on brittle materials or microstructural constituents
Ultrasonic	136° diamond pyramid	15–50 μm (0.06–0.2 μin.)	4–18 μm (0.016–0.07 μin.)	800 gf	Direct readout onto meter or digital display	Surface better than 1.2 μm (0.004 μin.) for accurate work. Otherwise, up to 3 μm (0.012 μin.)	1200 (limited by speed at which operator can read display)	Thin stock and finished surfaces in any position	Calibration for Young's modulus necessary, 100% testing of finished parts. Completely nondestructive

Source: Ref 4

ened and quenched steel piece is tempered by reheating at a controlled and relatively low temperature and then cooled at a control rate and time, it is possible to produce a wide range of desired hardness levels. By using a hardness test to monitor the end results, the operator is able to determine and control the ideal temperatures and times so that a specified hardness may be obtained.

When large populations of materials make testing each workpiece impractical and a tighter control is demanded for a product, statistical process control (SPC) is usually incorporated. This means of statistical control can enable continual product manufacturing with minimum testing and a high level of quality. Because many hardness tests are done rapidly, they are well suited for use with SPC techniques. Users are cautioned that the proper testing procedures must be followed to ensure the high degree of accuracy necessary when using SPC.

Tensile Testing

The tensile test is the most common test used to evaluate the mechanical properties of materials. Tensile testing is normally conducted by the material producer and the results are supplied to the user as part of the material certification sheet. Since the tensile test is a destructive test, it is not performed directly on the supplied material. For wrought materials, the test specimens are taken from the same heat or lot of material that is supplied. In the case of castings, separate test bars are cast at the same time as the part casting and from the same material used to pour the part casting. Although the tensile test is not normally conducted by the user of the metal product, it is important for the user to understand the test and its results.

Unless the material specification requires an elevated temperature test, the tensile test is normally conducted at room temperature. Typical values reported on the material certification include the yield strength, the ultimate tensile strength, and the percent elongation. Since the modulus of elasticity is a structure insensitive property and not affected by processing, it is generally not required. The main advantages of the tensile test are, the stress state is well established, the test has been carefully standardized, and the test is relatively easy and inexpensive to perform.

The tensile properties of a material are determined by applying a tension load to a specimen and measuring the elongation or extension in a load frame such as the one shown in Fig. 5. The load can be converted to engineering stress *s* by dividing the load by the original cross-sectional area of the specimen. The engineering strain (*e*) can be calculated by dividing the change in gage length by the original gage length.

A typical stress-strain curve for a metal is shown in Fig. 6. The shape and magnitude of the stress-strain curve of a metal depends on its composition, heat treatment, prior history of plastic deformation, and the strain

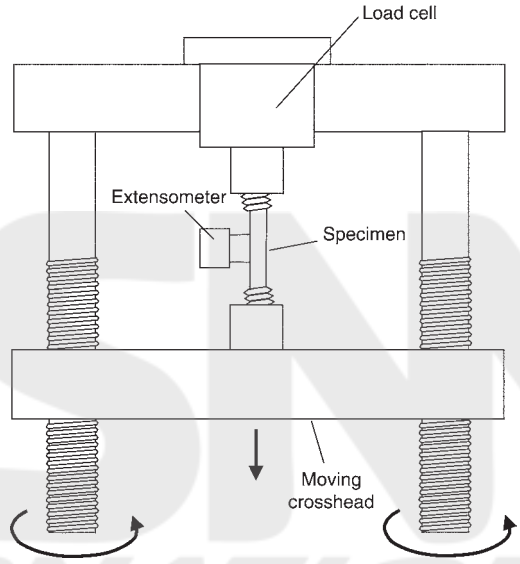


Fig. 5 Typical tensile test set-up. Source: Ref 3

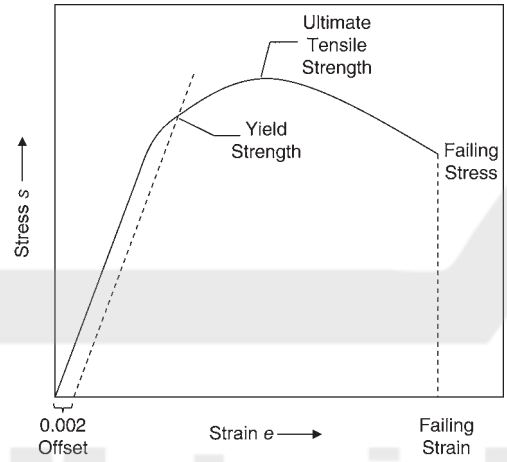


Fig. 6 Typical stress-strain curve. Source: Ref 3

rate, temperature, and state of stress imposed during the testing. The parameters used to describe the stress-strain curve of a metal are, the tensile strength, yield strength or yield point, percent elongation, and reduction in area. The first two are strength parameters and the last two are indications of ductility.

The yield strength (YS) is the stress required to produce a small specified amount of plastic deformation. The usual definition of this property is the offset yield strength determined by the stress corresponding to the intersection of the stress-strain curve offset by a specified strain. For metals without a definite yield point, the yield strength is determined by drawing

a straight line parallel to the initial straight line portion of the stress-strain curve. The line is normally offset by a strain of 0.2% (0.002).

As shown in Fig. 6, the ultimate tensile strength (UTS) is the maximum stress that occurs during the test. Although the tensile strength is the value most often listed from the results of tensile testing, it is not generally the value that is used in design. Static design of ductile metals is usually based on the yield strength, since most designs do not allow any plastic deformation. However, for brittle metals that do not display any appreciable plastic deformation, tensile strength is a valid design criterion.

Measures of ductility that are obtained from the tension test are the engineering strain at fracture (e_f) and the reduction of area at fracture (q). Both are usually expressed as percentages, with the engineering strain at failure often reported as the percent elongation.

Chemical Analysis

The overall chemical composition of metals and alloys is most commonly determined by x-ray fluorescence (XRF) and optical emission spectroscopy (OES). While these methods work well for most elements, they are not useful for dissolved gases and some nonmetallic elements that can be present in metals as alloying or impurity elements. High temperature combustion and inert gas fusion methods are typically used to analyze for dissolved gases (oxygen, nitrogen, hydrogen) and, in some cases, carbon and sulfur in metals.

A number of methods can be used to obtain information about the chemistry of the first one to several atomic layers of samples of metals, as well as of other materials, such as semiconductors and various types of thin films. Of these methods, the scanning Auger microprobe (SAM) is the most widely used.

Metallography

Metallography is the scientific discipline of examining and determining the constitution and the underlying structure of the constituents in metals and alloys. The objective of metallography is to accurately reveal material structure at the surface of a sample and/or from a cross-section specimen. For example, cross-sections cut from a component or sample may be macroscopically examined by light illumination in order to reveal various important macrostructural features (on the order of 1 mm to 1 m or 0.04 in. to 3 ft), such as the ones shown in Fig. 7 and listed here:

- Flow lines in wrought products
- Solidification structures in cast products
- Weld characteristics, including depth of penetration, fusion zone size and number of passes, size of heat affected zone, and type and density of weld imperfections

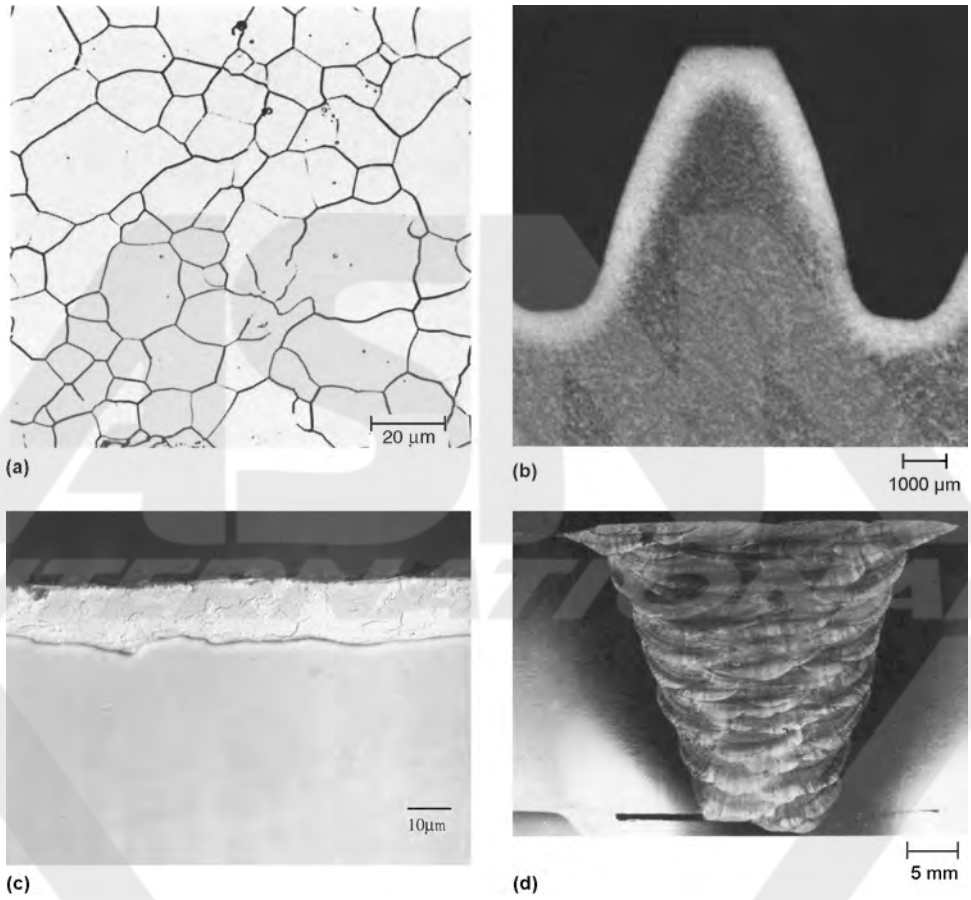


Fig. 7 Examples of uses for metallography. (a) Equiaxed ferrite grain size in plain carbon steel. (b) Ion carburized gear tooth showing case depth. (c) Microstructure of galvanized coating on steel, thickness and quality. (d) Multipass weld quality in type 304 stainless steel plate. Source: Ref 5, 6, and 7.

- General size and distribution of large inclusions and stringers
- Fabrication imperfections, such as laps, cold welds, folds, and seams, in wrought products
- Gas and shrinkage porosity in cast products
- Depth and uniformity of a hardened layer in a case hardened product

Macroscopic examination of a component surface is also essential in evaluating the condition of a material or the cause of failure. This may include:

- Characterization of the macrostructural features of fracture surfaces to identify fracture initiation site and changes in crack propagation process
- Estimations of surface roughness, grinding patterns, and honing angles
- Evaluation of coating integrity and uniformity

- Determination of extent and location of wear
- Estimation of plastic deformation associated with various mechanical processes
- Determination of the extent and form of corrosive attack; readily distinguishable types of attack include pitting, uniform, crevice, and erosion corrosion
- Evaluation of tendency for oxidation
- Association of failure with welds, solders, and other processing operations

This listing of macrostructural features in the characterization of metals, though incomplete, represents the wide variety of features that can be evaluated by light macroscopy.

Nondestructive Testing

Nondestructive testing (NDT) and inspection techniques are commonly used to detect and evaluate flaws (irregularities or discontinuities) or leaks in engineering systems. Of the many different NDT techniques used in industry, liquid penetrant and magnetic particle testing account for about one-half of all NDT, ultrasonics and x-ray methods about another third, eddy current testing about 10%, and all other methods for only about 2%. It should be noted that the techniques reviewed in this book are by no means all of the NDT techniques utilized. However, they do represent the most commonly employed methods. A simplified breakdown of the complexity and relative requirements of the five most frequently used NDT techniques is shown in Table 2, and the common NDT methods are com-

Table 2 The relative uses and merits of various nondestructive testing methods

	Test method				
	Ultrasonics	**X-ray**	**Eddy current**	**Magnetic particle**	**Liquid penetrant**
Capital cost	Medium to high	High	Low to medium	Medium	Low
Consumable cost	Very low	High	Low	Medium	Medium
Time of results	Immediate	Delayed	Immediate	Short delay	Short delay
Effect of geometry	Important	Important	Important	Not too important	Not too important
Access problems	Important	Important	Important	Important	Important
Type of defect	Internal	Most	External	External	Surface breaking
Relative sensitivity	High	Medium	High	Low	Low
Formal record	Expensive	Standard	Expensive	Unusual	Unusual
Operator skill	High	High	Medium	Low	Low
Operator training	Important	Important	Important	Important	
Training needs	High	High	Medium	Low	Low
Portability of equipment	High	Low	High to medium	High to medium	High
Dependent on material composition	Very	Quite	Very	Magnetic only	Little
Ability to automate	Good	Fair	Good	Fair	Fair
Capabilities	Thickness gaging: some composition testing	Thickness gaging	Thickness gaging; grade sorting	Defects only	Defects only

Source: Ref 8

pared in Table 3. Detailed information on the various types of NDT methods can be obtained from the American Society for Nondestructive Testing (Columbus, Ohio), *Nondestructive Testing*, Volume 03.03, published annually by ASTM (Philadelphia, Pennsylvania), and in *Nondestructive Evaluation and Quality Control*, Volume 17, *ASM Handbook*.

The terms *nondestructive testing* (NDT), and *nondestructive inspection*, (NDI), are considered synonymous. They both refer to a process or procedure, such as ultrasonic or radiographic inspection, for determining the quality or characteristics of a material, part, or assembly, without permanently altering the subject or its properties. All NDT or NDI methods are used to find internal anomalies or flaws in a structure without degrading its properties or impairing its serviceability. The term *flaw* is a general term that is used to imply any irregularity, imperfection, or discontinuity contained in a material, part, or assembly. A flaw that has been evaluated as rejectionable is usually termed a *defect*. The quantitative analysis of NDT/NDI findings to determine whether the material, part, or assembly will be acceptable for its function, despite the presence of flaws, is called

Table 3 Comparison of some nondestructive testing methods

Application	Characteristics detected	Advantages	Limitations	Example of use
Ultrasonics	Changes in acoustic impedance caused by cracks, nonbonds, inclusions, or interfaces	Can penetrate thick materials; excellent for crack detection; can be automated	Normally requires coupling to material either by contact to surface or immersion in a fluid such as water. Surface needs to be smooth.	Adhesive assemblies for bond integrity; laminations; hydrogen cracking
Radiography	Changes in density from voids, inclusions, material variations; placement of internal parts	Can be used to inspect wide range of materials and thicknesses; versatile; film provides record of inspection	Radiation safety requires precautions; expensive; detection of cracks can be difficult unless perpendicular to x-ray film.	Pipeline welds for penetration, inclusions, and voids; internal defects in castings
Visual optical	Surface characteristics such as finish, scratches, cracks, or color; strain in transparent materials; corrosion	Often convenient; can be automated	Can be applied only to surfaces, through surface openings, or to transparent material	Paper, wood, or metal for surface finish and uniformity
Eddy current	Changes in electrical conductivity caused by material variations, cracks, voids, or inclusions	Readily automated; moderate cost	Limited to electrically conducting materials; limited penetration depth	Heat exchanger tubes for wall thinning and cracks
Liquid penetrant	Surface openings due to cracks, porosity, seams, or folds	Inexpensive, easy to use, readily portable, sensitive to small surface flaws	Flaw must be open to surface. Not useful on porous materials or rough surfaces	Turbine blades for surface cracks or porosity; grinding cracks
Magnetic particles	Leakage magnetic flux caused by surface or near-surface cracks, voids, inclusions, or material or geometry changes	Inexpensive or moderate cost, sensitive both to surface and near- surface flaws	Limited to ferromagnetic material; surface preparation and post-inspection demagnetization may be required	Railroad wheels for cracks; large castings

Source: Ref 8

nondestructive evaluation (NDE). With NDE, a flaw can be classified by its size, shape, type, and location, allowing the investigator to determine whether or not the flaw(s) is acceptable. Damage tolerant design approaches are based on the philosophy of ensuring safe operation in the presence of flaws.

Liquid Penetrant Inspection

Liquid penetrant inspection is a nondestructive method used to find discontinuities that are open to the surface of solid, essentially nonporous materials. Indications of flaws can be found regardless of the size, configuration, internal structure, and chemical composition of the workpiece being inspected, as well as flaw orientation. Liquid penetrants can seep into (and be drawn into) various types of minute surface openings (as fine as 0.1 μm or 4 μin. in width) by capillary action, as illustrated in Fig. 8. Therefore, the process is well suited to detect all types of surface cracks, laps, porosity, shrinkage areas, laminations, and similar discontinuities. It is used extensively to inspect ferrous and nonferrous metal wrought and cast products, powder metallurgy parts, ceramics, plastics, and glass objects.

The liquid penetrant inspection method is relatively simple to perform, there are few limitations due to specimen material and geometry, and it is inexpensive. The equipment is very simple, and the inspection can be performed at many stages in the production of the part, as well as after the part is placed in service. Relatively little specialized training is required to perform the inspection. In some instances, liquid penetrant sensitivity is greater for ferromagnetic steels than that of magnetic particle inspection.

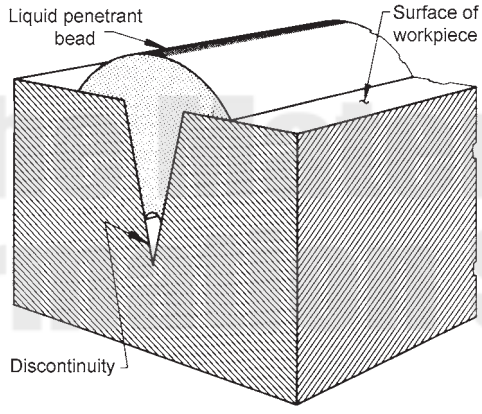


Fig. 8 Bead of liquid penetrant formed when, after excess penetrant has been removed from a workpiece surface, the penetrant remaining in a discontinuity emerges to the surface until an equilibrium is established. Source: Ref 9

The major limitation of liquid penetrant inspection is that it can detect only imperfections that are open to the surface; some other method must be used to detect subsurface defects and discontinuities. Another factor that can inhibit the effectiveness of liquid penetrant inspection is the surface roughness of the object. Extremely rough and porous surfaces are likely to produce false indications.

Magnetic Particle Inspection

Magnetic particle inspection is used to locate surface and subsurface discontinuities in ferromagnetic materials. The method is based on the fact that when a material or part being tested is magnetized, discontinuities that lie in a direction generally transverse to the direction of the magnetic field cause a leakage field to form at and above the surface of the part. The presence of the leakage field; and, therefore, the presence of the discontinuity, is detected by the use of finely divided ferromagnetic particles applied over the surface. Some of the particles are gathered and held by the leakage field. The magnetically held particles form an outline of the discontinuity and generally indicate its location, size, shape, and extent. Magnetic particles are applied over a surface either as dry particles or as wet particles in a liquid carrier such as water and oil. Different types of arrangements and coils can be used to control the direction of the magnetic field, as shown for the cases of circular and longitudinal magnetization in Fig 9.

Nonferromagnetic materials cannot be inspected by this method. Such materials include aluminum alloys, magnesium alloys, copper and copper alloys, lead, titanium and titanium alloys, and austenitic stainless steels.

The principal industrial uses of magnetic particle inspection are final inspection, receiving inspection, in-process inspection and quality control,

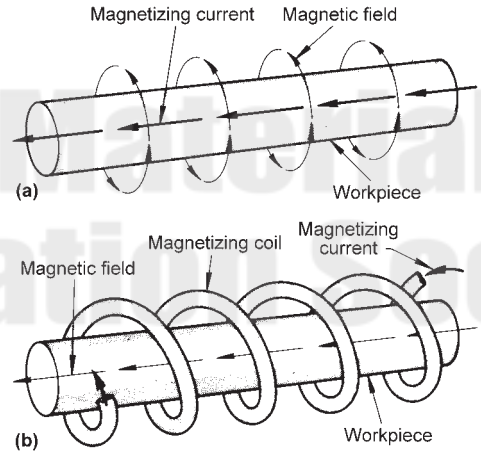


Fig. 9 Magnetized bars showing directions of magnetic field. (a) Circular. (b) Longitudinal. Source: Ref 10

maintenance and overhaul in the transportation industries, plant and machinery maintenance, and, inspection of large components.

Although in-process magnetic particle inspection is used to detect discontinuities and imperfections in material and parts as early as possible in the sequence of operations, final inspection is required to ensure that rejectable discontinuities and imperfections detrimental to part use and function have not developed during processing.

Eddy Current Inspection

Eddy current inspection is based on the principles of electromagnetic induction and is used to identify or differentiate a wide variety of physical, structural, and metallurgical conditions in electrically conductive ferromagnetic and nonferromagnetic metals and metal parts. The part to be inspected is placed within or adjacent to an electric coil in which an alternating current is flowing. As shown in Fig. 10, this alternating current, called the exciting current, causes eddy currents to flow in the part as a result of electromagnetic induction.

Eddy current inspection is used:

- To measure and identify conditions and properties related to electrical conductivity, magnetic permeability, and physical dimensions (primary factors affecting eddy current response)
- To detect seams, laps, cracks, voids, and inclusions
- To sort dissimilar metals and detect differences in their composition,

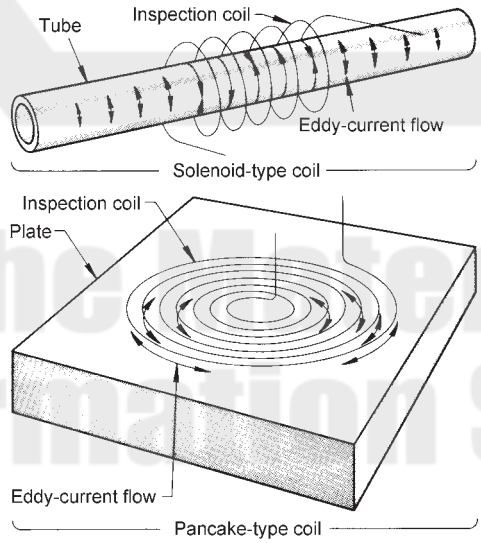


Fig. 10 Two common types of inspection coils and the patterns of eddy current flow generated by the exciting current in the coils. Solenoid type coil is applied to cylindrical or tubular parts; pancake type coil, to a flat surface. Source: Ref 11

microstructure, and other properties, such as grain size, heat treatment, and hardness
- To measure the thickness of a nonconductive coating on a conductive metal, or the thickness of a nonmagnetic metal coating on a magnetic metal

Because eddy current inspection is an electromagnetic induction technique, it does not require direct electrical contact with the part being inspected. Eddy current is adaptable to high speed inspection, and because it is nondestructive, it can be used to inspect an entire production output if desired. The method is based on indirect measurement, and the correlation between instrument readings and the structural characteristics and serviceability of parts being inspected must be carefully and repeatedly established.

Radiographic Inspection

Three basic elements of radiography include a radiation source, the testpiece or object being evaluated, and a sensing material. These elements are shown schematically in Fig. 11. Radiography is based on differential absorption of penetrating radiation−either electromagnetic radiation of very short wavelength or particulate radiation−by the part or test piece (object) being inspected. Because of differences in density and variations in thickness of the part, or differences in absorption characteristics caused by variations in composition, different portions of a testpiece absorb different amounts of penetrating radiation. Unabsorbed radiation passing

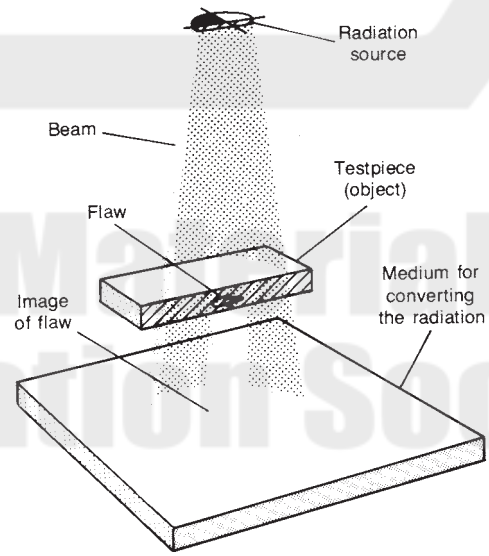


Fig. 11 Schematic of the basic elements of a radiographic system showing the method of sensing the image of an internal flaw in a plate of uniform thickness. Source: Ref 12

through the part can be recorded on film or photosensitive paper, viewed on a fluorescent screen, or monitored by various types of radiation detectors. The term *radiography* usually implies a radiographic process that produces a permanent image on film (conventional radiography) or paper (paper radiography or xeroradiography), although, in a broad sense, it refers to all forms of radiographic inspection. When inspection involves viewing of a real-time image on a fluorescent screen or image intensifier, the radiographic process is termed *real-time inspection*. When electronic, nonimaging instruments are used to measure the intensity of radiation, the process is termed *radiation gaging*. Tomography, a radiation inspection method adapted from the medical computerized axial tomography CAT scanner, provides a cross-sectional view of an inspection object. All the previous terms are mainly used in connection with inspection that involves penetrating electromagnetic radiation in the form of x-rays or gamma rays. *Neutron radiography* refers to radiographic inspection using neutrons rather than electromagnetic radiation.

In conventional radiography, an object is placed in a beam of x-rays and the portion of the radiation that is not absorbed by the object impinges on a detector such as film. The unabsorbed radiation exposes the film emulsion, similar to the way that light exposes film in photography. Development of the film produces an image that is a twodimensional *shadow picture* of the object. Variations in density, thickness, and composition of the object being inspected cause variations in the intensity of the unabsorbed radiation and appear as variations in photographic density (shades of gray) in the developed film. Evaluation of the radiograph is based on a comparison of the differences in photographic density with known characteristics of the object itself or with standards derived from radiographs of similar objects of acceptable quality.

Radiography is used to detect features of a component or assembly that exhibit differences in thickness or physical density compared with surrounding material. Large differences are more easily detected than small ones. In general, radiography can detect only those features that have a reasonable thickness or radiation path length in a direction parallel to the radiation beam. This means that the ability of the process to detect planar discontinuities such as cracks depends on proper orientation of the testpiece during inspection. Discontinuities such as voids and inclusions, which have measurable thickness in all directions, can be detected as long as they are not too small in relation to section thickness. In general, features that exhibit differences in absorption of a few percent compared with the surrounding material can be detected.

Ultrasonic Inspection

Ultrasonic inspection is a nondestructive method in which beams of high frequency acoustic energy are introduced into a material to detect

surface and subsurface flaws, to measure the thickness of the material, and to measure the distance to a flaw. An ultrasonic beam travels through a material until it strikes an interface or discontinuity such as a flaw. Interfaces and flaws interrupt the beam and reflect a portion of the incident acoustic energy. The amount of energy reflected is a function of (a) the nature and orientation of the interface or flaw; and, (b) the acoustic impedance of such a reflector. Energy reflected from various interfaces and flaws can be used to define the presence and locations of flaws, the thickness of the material, and the depth of a flaw beneath a surface. Pulse echo and through transmission, two types of ultrasonic inspection, are illustrated in Fig. 12.

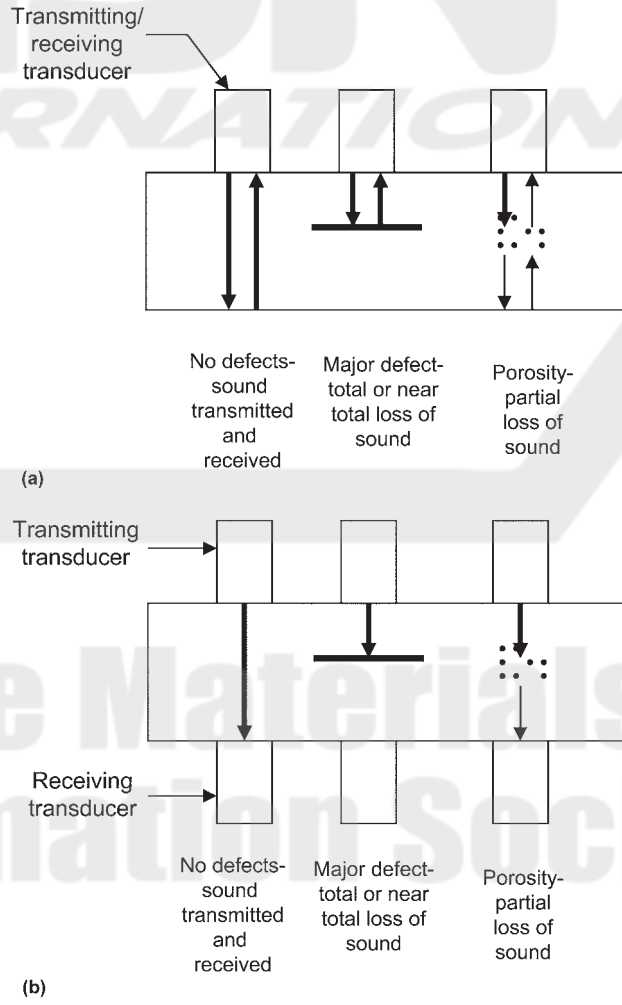


Fig. 12 Two types of ultrasonic inspection. (a) Pulse echo. (b) Through transmission

Most ultrasonic inspections are performed using a frequency between 1 and 25 MHz. Short shock bursts of ultrasonic energy are aimed into the material from the ultrasonic search unit of the ultrasonic flaw detector instrument. The electrical pulse from the flaw detector is converted into ultrasonic energy by a piezoelectric transducer element in the search unit. The beam pattern from the search unit is determined by the operating frequency and size of the transducer element. Ultrasonic energy travels through the material at a specific velocity that is dependent on the physical properties of the material and on the mode of propagation of the ultrasonic wave. The amount of energy reflected from or transmitted through an interface, other type of discontinuity, or reflector depends on the properties of the reflector. These phenomena provide the basis for establishing two of the most common measurement parameters used in ultrasonic inspection: the amplitude of the energy reflected from an interface or flaw; and, the time required (from pulse initiation) for the ultrasonic beam to reach the interface or flaw.

REFERENCES

1. S.M. Purdy, Macroetching, *Metallography and Microstructures*, Vol 9, *ASM Handbook*, ASM International, 2004, p 313–324
2. J.D. Meyer, Machine Vision and Robotic Inspection Systems, *Nondestructive Evaluation and Quality Control*, Vol 17, *ASM Handbook*, ASM International, 1992, p 29–45
3. F.C. Campbell, *Elements of Metallurgy and Engineering Alloys*, ASM International, 2008
4. A. Fee, Selection and Industrial Applications of Hardness Tests, *Mechanical Testing and Evaluation*, Vol 8, *ASM Handbook*, ASM International, 2000, p 260–277
5. A. O. Benscoter and B.L. Bramfitt, Metallography and Microstructures of Low-Carbon and Coated Steels, *Metallography and Microstructures*, Vol 9, *ASM Handbook*, ASM International, 2004, p 588–607
6. Metallography and Microstructures of Case-Hardening Steel, *Metallography and Microstructures*, Vol 9, *ASM Handbook*, ASM International, 2004, p 627–643
7. Metallography and Microstructures of Weldments, *Metallography and Microstructures*, Vol 9, *ASM Handbook*, ASM International, 2004, p 1047–1056
8. L. Cartz, Quality Control and NDT, *Nondestructive Testing*, ASM International, 1995, p 1–13
9. J.S. Borucki and G. Jordan, Liquid Penetrant Inspection, *Nondestructive Evaluation and Quality Control*, Vol 17, ASM Handbook, 1989, p 491–511
10. A. Lindgren, Magnetic Particle Inspection, *Nondestructive Eval-*

uation and Quality Control, Vol 17, ASM Handbook, 1989, p 89–128
11. Eddy Current Inspection, *Nondestructive Evaluation and Quality Control*, Vol 17, ASM Handbook, 1989, p 164–194
12. Radiographic Inspection, *Nondestructive Evaluation and Quality Control*, Vol 17, ASM Handbook, 1989, p 295–357

CHAPTER 2

Visual Inspection

VISUAL INSPECTION is perhaps the most important method of inspection of materials. *Visual inspection* is defined as the examination using the naked eye, alone or in conjunction with various magnifying devices, without changing, altering, or destroying the material involved.

To do a good job of visual inspection requires some knowledge of what you are looking at. It is good to have as much knowledge as possible of the product being examined. You must not only discover defects, but also be able to evaluate them from the point of usefulness or rejection. Knowledge of the cause of defective materials helps in future prevention. You should know how it may be abused. You should also be familiar with the types of defects that normally might be encountered in such a part, e.g., scabs, seams, and laminations in steel mill products; and corrosion, erosion, and physical abuse on parts that have been in service.

Visual Inspection Procedure

The part should first be carefully examined with the naked eye. Then, magnifying devices may be used to further examine suspect areas revealed by the naked eye examination.

Try to account for all unusual surface markings and conditions. Think in terms of depth effect and sharpness of penetration (stress concentration). Note any discoloration and determine the cause (heat? corrosion?).

Markings. Observe any identification markings. Such markings may identify the manufacturer, date of manufacture, original size of material, material specification, and number of the original heat of steel traceable to analysis and physical properties. If you are unfamiliar with identification marking procedures, do not assume a part is unmarked. Find out by asking manufacturers how they identify such products.

These are obvious markings. But there are also many hidden markings used by manufacturers, which they often do not readily reveal. One example is the wire rope industry. Some hemp center wire ropes have a single fiber wrapped with the group that reveals the name of the manufacturer and also the type of wire rope. It is difficult to distinguish this fiber and requires some care to separate and flatten it; sometimes soaking in hot water helps. Nearly all wire ropes have an identifying strand. It may be colored plastic, or the diameter of a single wire, or part of the construction configuration. Bolt heads are another good example of markings which give information. In addition to the radial markings on the heads which tell the strength level, the manufacturer's markings are often present.

Abuse. Look for evidence of abuse. In the case of failure, try to decide if the abuse occurred before or after the failure. Failures often involve severe trauma. Parts flying about after the fact can suffer severe abuse that may be misinterpreted as being related to the cause of the failure. If a part is distorted, try to ascertain the type, direction, and intensity of the load necessary to produce this distortion.

Heat Effects. Defects are often caused by heating problems. These heating problems generally leave telltale indications. The indications are the result of the complex oxide systems of iron and other alloying elements. At lower temperatures, steel parts with a relatively bright surface may display characteristic temper colors. These colors vary somewhat with the composition and heat treatment. A variety of colors are produced within the temperature range of 195 to 370 °C (380 to 700 °F). Charts can be consulted for the approximate temperature reached (Ref 1).

As the temperature rises, a heavier scale of a different type forms. The formation temperature varies with the type of alloy. It should be noted that a scale layer may be quite thick. This does not necessarily indicate a great metal loss. Generally, the ratio of lost metal to oxygen in a scale layer is 8:1. This means a 3 mm (⅛ in.) thick scale layer represents only about 0.4 mm (0.015 in.) of lost metal. Since scaling can occur at such a low temperature, the microstructure may not have been altered appreciably (assuming the lower critical temperature to be 723 °C, or 1333 °F).

The color of heat scale should also be observed. The brown- to red-colored scale layers on steel indicate the more completely oxidized iron oxide (Fe_2O_3). The black, mill-scale type of oxides are composed of the incompletely oxidized form of iron, black iron oxide (Fe_3O_4). This indicates an oxygen deficiency at the time of heating and may mean much higher temperatures of formation. There are other temperature indications associated with the observation of materials while being exposed to heat. Such materials just begin to glow dull red (in a dark area) at 620 to 650 °C (1150 to 1200 °F). The colors change as the temperature increases.

Corrosion Scaling. Scaling of materials is not necessarily associated with just heat. Corrosion also may create scale-type deposits. In some cases, these may be indistinguishable from heat scale. Often, valuable

data can be gained from scale analysis. It may be a good idea to remove scrapings and label for future tests. Try not to mix grease, paint, coating material, and mill scale with the corrosion product. If parts have internal and external surfaces, do not mix outside scrapings with inside scrapings. If corrosion is suspected, determine if it is localized or general. Is it uniform or selective (pitting)? Is it in an area of contact with other materials? Does it have special characteristics indicative of high velocity flow (e.g., erosion, cavitation)? Does it leave an unusual appearing corrosion product?

Cracking. If cracking is noted during visual examination, it is important to characterize the cracking. Is the crack straight or does it follow an irregular path? Is there one crack or a series of cracks? Are the cracks open or tightly closed? Are the cracks associated with markings of any kind? Are the cracks located in areas of natural stress concentration? Are the cracks associated with welding, for either fabrication or repair? Are the cracks associated in any way with evidence of corrosion, e.g., corrosion product, corrosion pitting?

If cracking is observed during the inspection of raw materials or in materials in process of manufacture or finished products, it should be explored for the purpose of determining acceptance or rejection. This is usually done by grinding to determine depth and extent. In such grinding, care must be exercised so that heat and stress do not cause extension of cracking. Grinding, if not done with care, can also close tight cracks by causing the adjacent metal to flow over the crack. It is often necessary to use methods such as magnetic particle testing or dye penetrant inspection to be assured that cracking is completely removed.

Once the crack has been completely removed, judgment must be exercised as to whether the part (a) can be used as is, (b) can be repaired, or (c) must be rejected and scrapped. Some specifications prescribe how much cracking is acceptable. The American Petroleum Institute, in many of its specifications, allows surface defects if the depth is no greater than 12½ percent of the wall thickness. Where no such specification exists, the ultimate usage must be considered.

Of primary consideration is dynamic loading versus static loading. Other helpful considerations in making an acceptance or rejection are drawing tolerances for the part and design information such as safety factors.

Once a decision is made, it is necessary to determine if welding repair is necessary or if the part can be used as is, with the defects ground out. Welds may cause a new set of problems. If the weld can be avoided, feather the ground-out area to a generous radius to avoid stress concentrations. The smoother this surface is, the less likely it is to crack again.

Straight, sharp, open, single cracks (Fig. 1a) are usually associated with very high stresses and/or material of lowered ductility. They may be associated with suddenly applied (impact) loads. Open cracks may also be

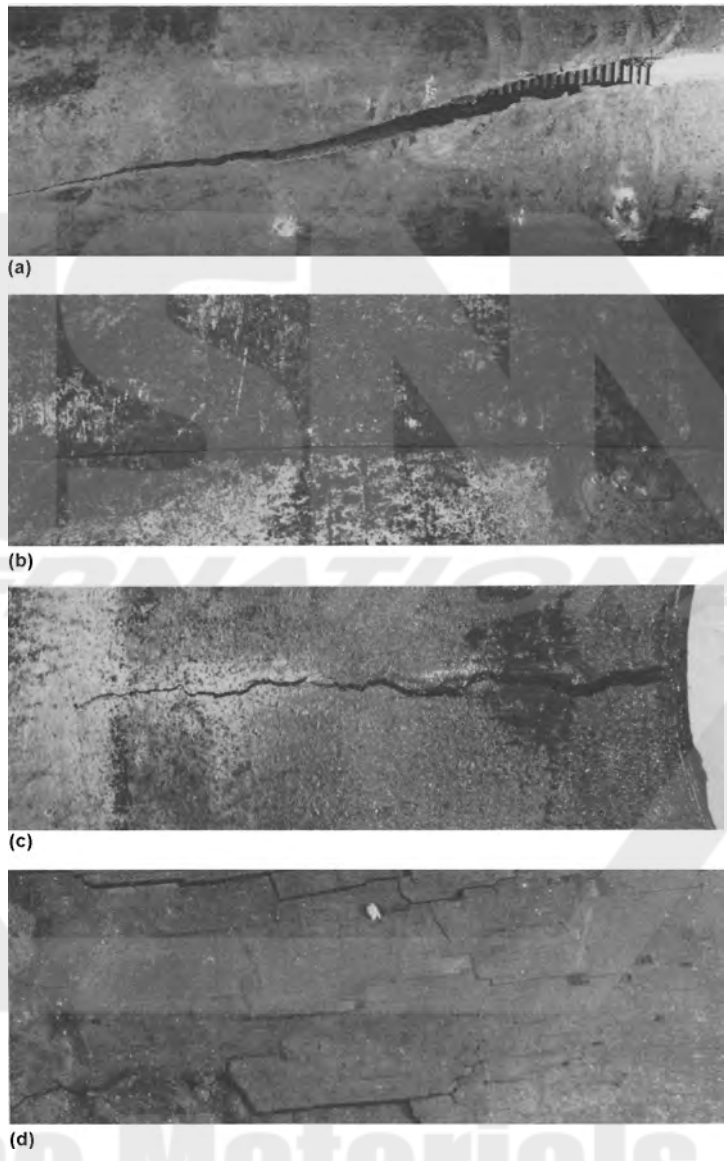


Fig. 1 Four types of cracks: (a) sharp and open, (b) sharp and tight, (c) jagged, (d) multiple. Source: Ref 1

associated with locked-up internal stresses of a large magnitude. As an example, parts with high locked-up stresses may spring open when cut with a saw. Conversely, when locked-up stresses act in the opposite direction the cracks will be tight, due to compressive stresses (Fig. 1b), and the part might clamp shut on the saw blade when cutting. Cracks which are jagged in configuration (Fig. 1c) may be indicative of ductile tearing or may represent a separation at the grain boundaries (intergranular). Irregular-shaped cracking is usually indicative of slower crack propagation. Multi-

ple cracks (Fig. 1d) are often associated with corrosion, stress corrosion, corrosion fatigue, thermal fatigue, or localized trauma.

It is important to note the location and orientation of cracks. Cracks are related to high stresses of either internal or external origin. Cracks will occur at the weakest point, usually a point of stress concentration.

Stress concentration is the nonuniform distribution of stresses in a loaded part. Stress concentrations take various forms (Fig. 2). They may occur as changes in section size. They may be in the form of a radius which can be gentle and smooth (good) or sharp and rough (bad). Purposely made markings should be suspect. These can take the form of rough machine grooves, coarse grinding marks, stamped numbers and letters, indentations (Fig. 3), holes, keyways, and scratches.

Another source of crack initiation is welding. Particular attention should be focused on welds, where cracks are in the general area, whether these are fabrication or repair welds. Look at the size (small welds are undesirable); look at the appearance (good welds look well made). Look for arc strikes and splatter (bad). Look for undercutting, which produces a stress concentration groove adjacent to the weld.

Measurement is considered part of visual examination. Acceptance and rejection of materials may depend entirely on dimensions. Failure may also relate to dimensions. Most manufactured parts are made to some

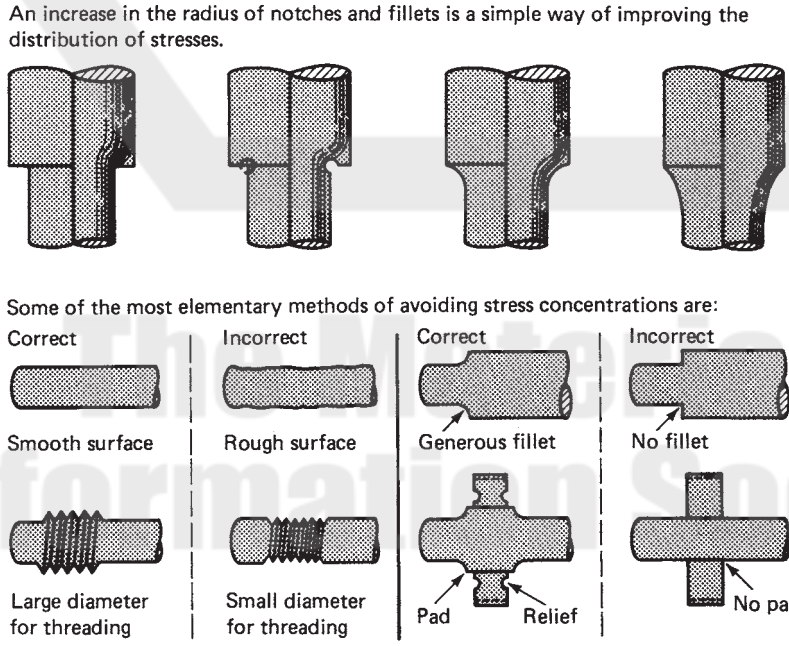


Fig. 2 Examples of stress concentration problems and solutions. Top, courtesy of Battle Memorial Institute; bottom, courtesy of McQuaid. Source: Ref 1

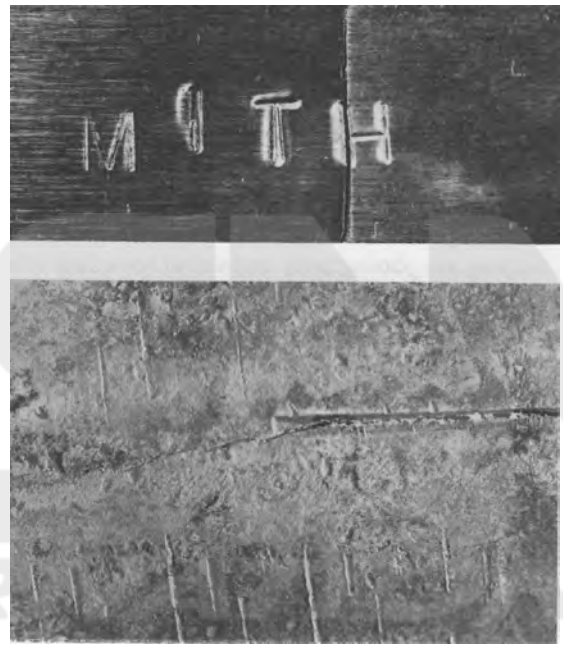


Fig. 3 Two examples of cracking caused by indentations that produced stress concentrations. Source: Ref 1

tolerance of measurement. Some of these may be to plus or minus several inches or even larger, while others may be as little as one ten-thousandth of an inch. Most are somewhere in between. If drawings and specifications are available, critical dimensions and tolerances can be readily determined.

Measuring involves overall dimensions, inside and outside diameters, depths of holes, radii, thread sizes, and surface finishes. Depths of pits and lengths of cracks can also be measured if present. Most such measurements are simple, but some require very specialized equipment.

Results and Record Keeping. There must be a way to communicate the results of the visual examination, both to others and for future reference. Written notes describing what has been observed are the most widely used method of record keeping. Printed forms with blanks to be filled in may aid the note-taking procedure. Dimensional measurements can be in the form of notes or as designations on sketches. Other common methods of record keeping are photography and verbal recording using a tape recorder. A good, clear, enlarged photograph is worth many words. Color photography should be considered. Not so commonly used are motion pictures and video tapes.

When recording results of visual examination, try to describe the part as if the reader had never seen it. Start by generally explaining what was examined, gradually becoming more specific. Describe the part and its condition thoroughly. This will assure good communication and lend credibil-

ity to your account. It will also serve as a good refresher if you are asked to explain your findings at a later date.

Visual Inspection Tools

Tools for visual inspection can be grouped into six categories:

- Magnifying devices
- Lighting for visual inspection
- Measuring devices
- Miscellaneous measuring devices
- Record-keeping devices
- Macroetching

Magnifying Devices

Magnifiers can be characterized by magnifying power, focal length, and lens type.

Magnifying Power. An object appears to increase in size as it is brought closer to the eye. In determining magnifying power, the true size of the object is what the image appears to the eye at 10 in. (25 cm). The 10-inch value is used as a standard because this is the distance from the eye one usually holds a small object when examining it. Linear magnification is expressed in diameters. The letter × is normally used to designate the magnifying power of a lens, e.g., 10×.

If one could focus on an object at one inch (2.5 cm), it would appear 10 times larger. Since one cannot effectively focus the eye at one inch, a lens may be used to do so. Thus, magnification can be defined as the ratio of the apparent size of an object seen through a magnifier (known as a virtual image) to the size of the object as it appears to the unaided eye at 10 inches.

Focal Length. The focal length is the distance from the lens to the point at which parallel rays of light striking one side of a positive lens will be brought into focus on the opposite side. For lenses of short focal length, such as discussed here, light 30 to 40 feet, or 9 to 10 meters away can be considered parallel. The focal length can be determined by holding a lens such that light coming through a window, for example, will allow the image of the window or other object to focus sharply on a sheet of paper held behind the lens. The distance from lens to paper will then be the focal length. Once the focal length is known, the magnification of the lens can be determined, and vice versa. The shorter the focal length, the greater the magnifying power. The distance of the eye from the lens must be the same as the focal length. A lens with a one-inch focal length, for example, will have a magnifying power of 10 (10×). This is true if the lens is held one inch from an object and the eye is placed one inch from the lens.

In summary, the following formula determines magnifying power:

$$\text{Magnifying power (any positive lens)} = \frac{10}{\text{focal length}} \text{(in inches)}$$

With a simple method of determining focal length, it becomes easy to determine magnification.

Lens Types

All lenses are either convex (bulged out), concave (sunken in), or flat. More often they are a combination of these. The most common type found in the laboratory is the double convex lens. Lenses with one side convex and the other flat (plano-convex) are used in projectors and microscopes. All other magnifiers are lenses used in combination.

The degree of correction dictates the quality of the lens. Three inherent faults in lenses–all of which are correctable—are:

- *Distortion.* The image appears unnatural. The quality of the lens material and the grinding and polishing are both the causes of and the means for correcting this problem.
- *Spherical Aberration.* Light rays passing through the center of the lens and at the outer edges come to a focus at different points. (The distortion is worse on large diameter lenses than small.) Spherical aberration can be corrected by slight modification of the curved surfaces.
- *Chromatic Aberration.* This is a prism effect: when broken down into colors, the light rays do not focus at the same place. This may occur both as a lateral and as a longitudinal effect. It is correctable by use of compound lenses of different types of glass.

Below about five magnifications, one double convex lens is satisfactory. Two combined lenses will have a shorter focal length than either lens used alone. Higher magnification in simple magnifiers usually employs two or three lenses in combination. Twenty magnifications is about the maximum for these simple devices. A 20× magnifier will have a focal length and a field of view of about 6 mm (¼ in.).

When choosing a simple magnifier, consider the following corrected lenses for quality (Fig. 4):

- *Coddington Magnifier* (Fig. 4a). This device uses a double convex lens with a groove ground in the middle. This diaphragm-like groove improves image quality by eliminating marginal rays of light.
- *Double Plano-Convex Magnifier* (Fig. 4b). This two-lens magnifier gives partial chromatic correction and flatter field of view.
- *Hastings Triplet Magnifier* (Fig. 4c). This is a multiglass lens corrected for both spherical and chromatic aberration. This is the best of all hand-held magnifiers.

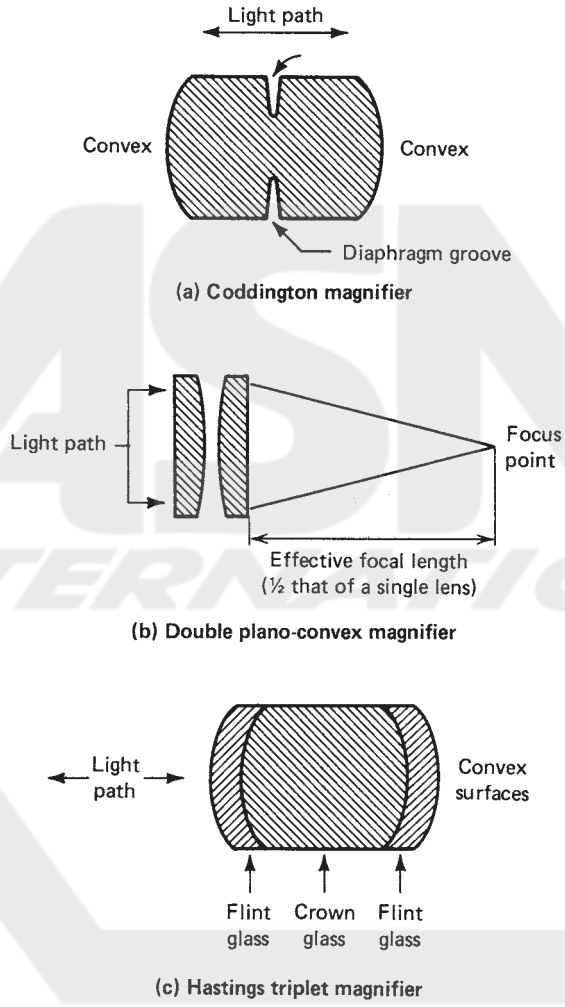


Fig. 4 Diagram showing lens corrections available in simple magnifiers. Source: Ref 1

The principal limiting factor for magnifying devices is depth of field. As magnification increases, the distance between the peaks and valleys (of an irregular surface) that are simultaneously in focus lessens. For example, at 100 magnifications the surface examined must be flat and polished. A variation of only $\frac{1}{1000}$ of an inch can be out of sharp focus.

In summary, as the magnifying power of a lens system increases: (a) there are fewer peaks and valleys in focus at the same time, (b) the area observable is smaller, and (c) the distance from lens to subject becomes shorter (in addition, among other problems, lighting the object is difficult). Common magnifiers rated over 20×, while readily available, are not very practical.

One other limiting factor in magnifying devices is light loss due to reflection. Lens surfaces can be coated with special antireflection coatings to reduce light loss, which may be particularly useful when the level of light is low.

Simple Magnifiers

Simple magnifiers come in many varieties, and new devices are regularly being developed. The following is an effort to group the various devices into categories:

- Hand-held lenses, single and multiple
- Pocket microscopes
- Self-supporting magnifiers
- Magnifying devices which can be worn attached to the head or in some manner be used like eyeglasses or in conjunction with eyeglasses
- Magnifying devices with built-in light sources

Hand-Held Lenses. These are available as a lens by itself, a lens with a frame and handle, or a lens that folds out or slides out of its own case. The fold-out type may include one to four lenses that can be used alone or in conjunction with one another. The size generally varies from 13 to 150 mm (½ to 6 in.) in diameter. They are available with either glass or plastic lenses.

The plastic (generally acrylic) lenses are shatterproof, but scratch easily. They are not capable of producing the lens corrections and quality of the glass lenses. The best of these hand-held lenses are the Hastings Triplet, Coddington, and the Plano-Convex, in that order (see Fig. 4 and 5).

Pocket Microscopes. Another variety of the hand-held magnifier are pocket microscopes (Fig. 6). These are generally small diameter tubes,

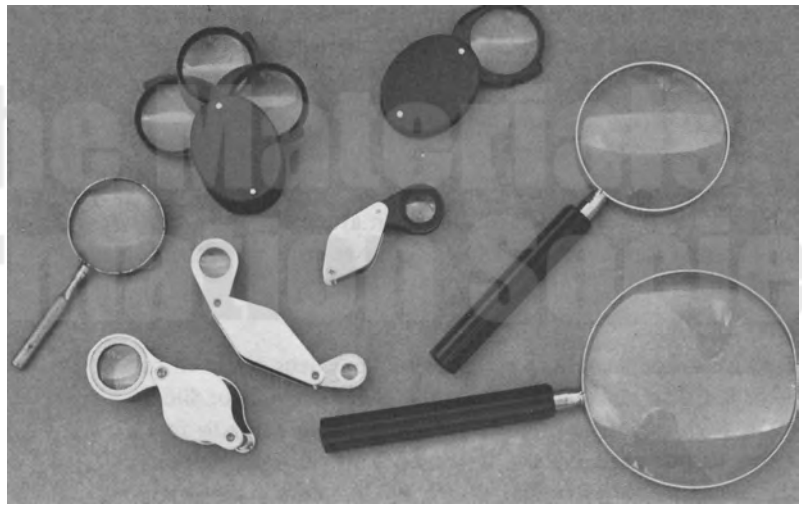


Fig. 5 Hand-held magnifiers. Source: Ref 1

about 13 mm (½ in.) in diameter and 150 mm (6 in.) in length, although they are also available in larger diameters. The smaller varieties are usually offered with magnification ranges of 25× to 60×. The subject end is cut at an angle or is somehow opened to allow maximum available light along with support. At these magnifications, the field of view and focal length are extremely limited, as is the available light. Auxiliary light is often a necessity. The larger-diameter units have lower magnifying power.

Self-Supporting Magnifiers. Self-supporting magnifiers (Fig. 7) are much like the hand-held magnifiers, except they free the hands to manipu-

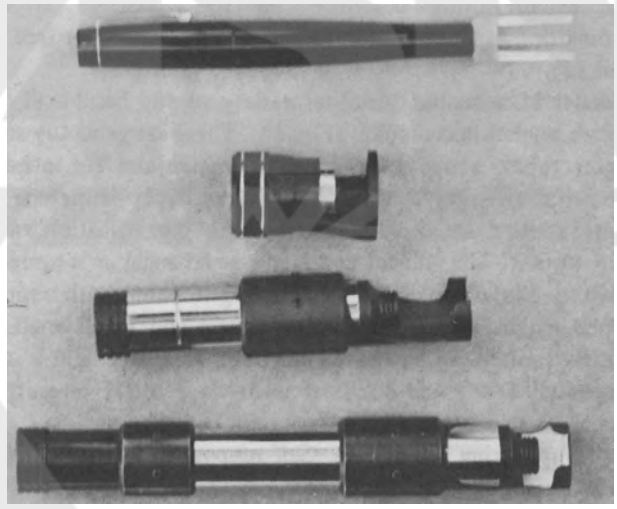


Fig. 6 Pocket microscopes. Source: Ref 1

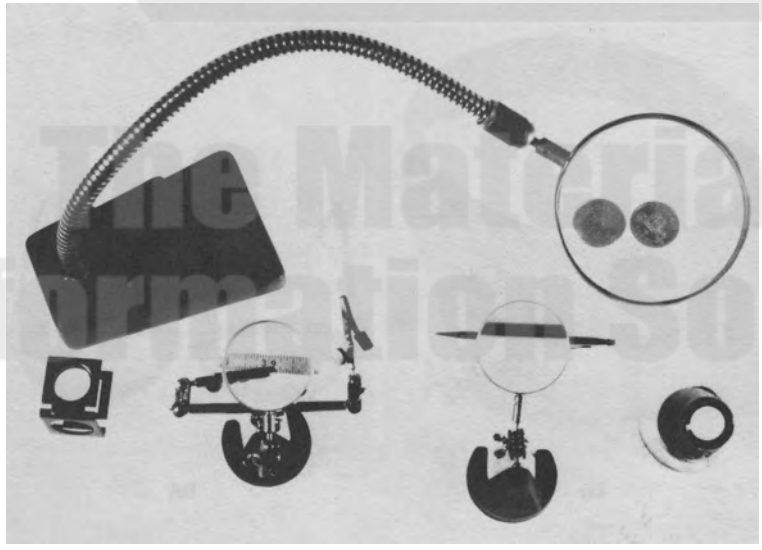


Fig. 7 Self-supporting magnifiers. Source: Ref 1

late the object being observed. They are generally low-power magnifying devices like the hand-held lenses. They are available as lenses with heavy bases and movable extension arms, lenses that sit directly on top of the object being viewed, and lenses that hang around the neck.

Magnifying Devices Which Are Eye Attachments. These magnifying devices are of two types. The visor type (Fig. 8a) has an adjustable band that fits over the head. This band supports a lens holder that tilts up and down for use when needed. The lens system may be two separate lenses or a continuous strip lens. It is also available with a loupe accessory for additional magnification. These visors may be worn with or without eyeglasses. Magnification offered is generally low (1½× to 3½×), but can be as high as 10× to 15×. They make excellent visual examination devices

Fig. 8 Magnifying devices that attach to the head or eye: (a) visor, (b) eyeglass loupe, (c) loupe. Source: Ref 1

because they can be comfortably worn for long periods of time and can be quickly tilted in place for use when needed.

The second type is the loupe. Loupes used without glasses (Fig. 8c) either can be held in the eye by use of eye muscles, like a monocle, or are available with a spring clip which wraps around the head. Loupes are also available that attach to eyeglasses as single or multiple lenses (Fig. 8b). These can be tilted in or out of use easily. The magnification range for such loupes is 2× to 18×.

Illuminated Magnifiers. Most of the magnifying devices described are also available with built-in light sources. To see details, good lighting is important. This is particularly true at the higher magnifications since the lens-to-subject distance is so short. Most light sources are either battery powered with flashlight batteries or equipped to plug into a standard wall outlet. The lights are usually incandescent, but are also available with fluorescent and ultraviolet light sources (Fig. 9).

Lighting

General Lighting. Very few indoor areas offer sufficient light to perform a proper visual examination. Sunlit areas are excellent for general examination, but may not be sufficient for examining internal areas such as bores and deep crevices. High-density fluorescent ceiling lighting offers good general-inspection lighting. For more specific overall lighting, there are three options (Fig. 10). One is a portable stand with an incandescent flood or spotlight bulb and reflector similar to those used by photog-

Fig. 9 Illuminated magnifiers. Source: Ref 1

Fig. 10 General lighting devices. Source: Ref 1

raphers. This gives a high-intensity source of light for a fairly large area. The stands are adjustable up and down, and the head swivels in all directions. This is a good light source for photographic recording. A word of caution on this type of light: bulb life is usually short (six hours), and considerable heat is generated.

When considering such equipment, it is wise to choose the sturdiest available. Two things to look for are heavy-duty swivel adjustments on the light head, and adequate cooling for the lamp base. These heavy-duty lights are available, but not as easy to find as the more common light-duty types. They are considerably more expensive, but easily worth the price.

The two other general lighting devices are swivel-arm incandescents and swivel-arm fluorescents. These come in a variety of shapes, sizes, intensities, and swivel-arm types. They provide less intensity and illuminate a smaller area than the flood or spot type described above. They are good for smaller areas and have longer lives. The fluorescent type has less intensity, but produces fewer shadows and is cooler operating. Many of the incandescent types have variable intensity controls. These lights can also be used in conjunction with magnifying devices.

Specific lighting devices are of high intensity and permit the light to be concentrated on a small spot. Several types are shown in Fig. 11. The more common varieties are incandescent. They usually utilize an adjust-

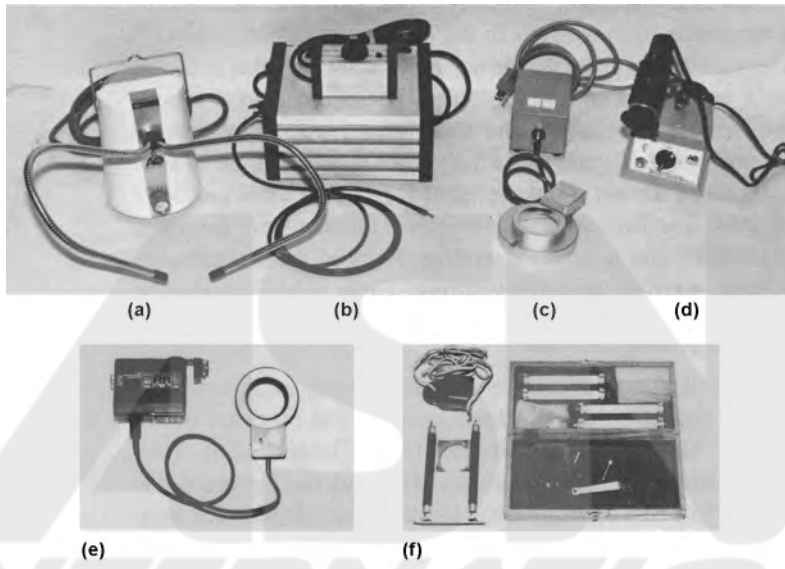


Fig. 11 Specific lighting devices: (a and b) fiber optics, (c) microscope ring light, (d) microscope light, (e) ring flash, (f) microscope illuminator using fluorescent and ultraviolet tubes. Source: Ref 1

able transformer and one or more diaphragms. They are on adjustable heads. These devices are most commonly sold as microscope lights. The problem with them is that they burn out and overheat easily, they do not have sufficient intensity, and they tend to produce an image of the light bulb filament on the subject being illuminated.

There are several other devices for high-intensity, highly localized lighting. Two of these are like the microscope lights previously described; one uses the halogen very-high-intensity light source, the other uses the carbon-arc light source. The latter offers the brightest light of all the available sources, but requires adjustments and arc replacement. A third available unit is a fiber optic device. This allows highly specific, high-intensity light to be brought very close to an object, even in confined quarters. It is excellent for higher-magnification viewing and extreme close-up photography.

Measuring Devices

Measuring devices are considered part of visual examination because they are used to record the results of the examination. Visual examination, among other purposes, includes checking to see if parts meet dimensional specifications. These devices are so numerous, including many which are highly specialized, that a separate volume could probably be written about them. Because of this, only those most commonly used will be mentioned.

Linear Measuring Devices. The most common measuring unit is the ruled straightedge. These come in many forms, such as the 12-inch rule and 36-inch yardstick, both of which are being replaced by meter sticks. Until the transition to the metric system is universal, both English and metric devices must be utilized. In addition, tape measures, which are available from 150 mm (6 in.) in length to over 30 mm (100 ft), are essential for visual examination.

Of the rules, the 150 and 300 mm (6 and 12 in.) steel rules are desirable. The 150 mm (6 in.) scales, some of which can be clipped to a shirt pocket, are available with several scales on each rule. Scales can be both English and metric on the same rule and may be subdivided to as small as $\frac{1}{100}$-inch divisions. These devices are also available with an adjustable 90° squaring edge to check for straightness.

Reticles. There are many magnifying devices available with built-in reticles (Fig. 12). Reticles made to measure nearly anything imaginable are available. If the desired reticle is not available, it can be custom-made.

Micrometers are extremely accurate mechanical devices. They are commonly used to measure to $\frac{1}{1000}$ of an inch and can be used to $\frac{1}{10,000}$ of an inch. Both inside and outside micrometers are available. Some incorporate both in the same unit. They are available with various measuring tips (these are normally hardened to prevent wear). Tips can be flat, rounded, pointed, or blade. Others may also be available or custom-made. Very little training is required to use these devices, but experience produces more consistently accurate results.

Optical comparators (Fig. 13) are excellent devices for both visual examination and measurement. A comparator produces a two-dimensional enlarged image of an object on a large ground-glass screen. It can be used with reflected light or background lighting (or a combination of both). Magnification is available from actual size to 50×. Comparison templates can be placed on the screen to check dimensional accuracy. Results can be readily photographed.

Miscellaneous Measuring Devices (Fig. 14). The dial indicator consists of a plunger-actuated dial, usually calibrated in $\frac{1}{1000}$ of an inch. It comes with a series of mechanical arms and clamping devices such that it can be attached to a fixed (rigid) object and reference measurements can be made. A common usage is attachment to a lathe bed to check both horizontal and circumferential dimensional variations.

There are many specialty measuring gages. Some of these are inside micrometers, tubing wall measuring micrometers (one rounded anvil and one flat anvil), depth gages, thread-measuring gages, protractors and bevel protractors (to measure angles), levels (to measure variation from horizontal), inside and outside calipers, hole and plug gages (to measure diameter and uniformity of holes), radius gages, screw-pitch gages, thickness gages (a series of leaves of various known thicknesses to check clearance).

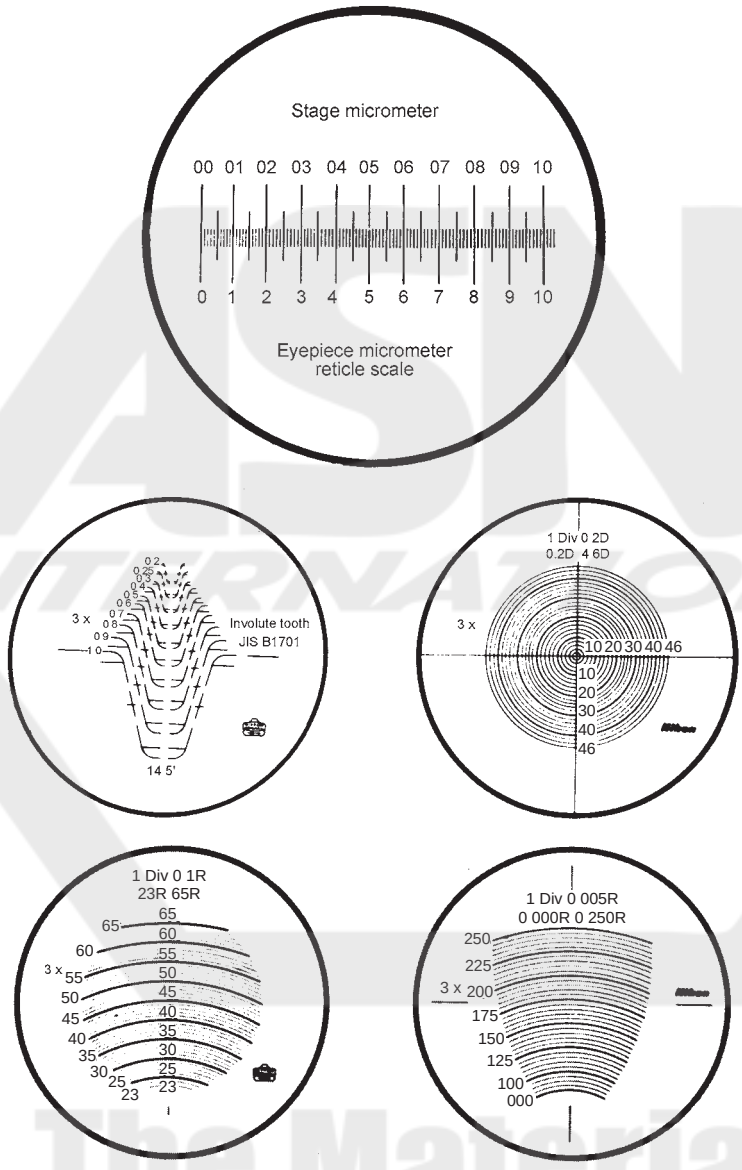


Fig. 12 Recticles. Source: Ref 1

Many such devices are specially designed and built for a particular application. A wide assortment are available as stock items, with many brands to choose from. Quality should not be sacrificed for cost on measuring devices. They should be kept in specially designed cases, be kept clean, and be lubricated as required. When making measurements, devices should not be forced or over tightened. Many of these tools come with calibration blocks and should be checked regularly for best results.

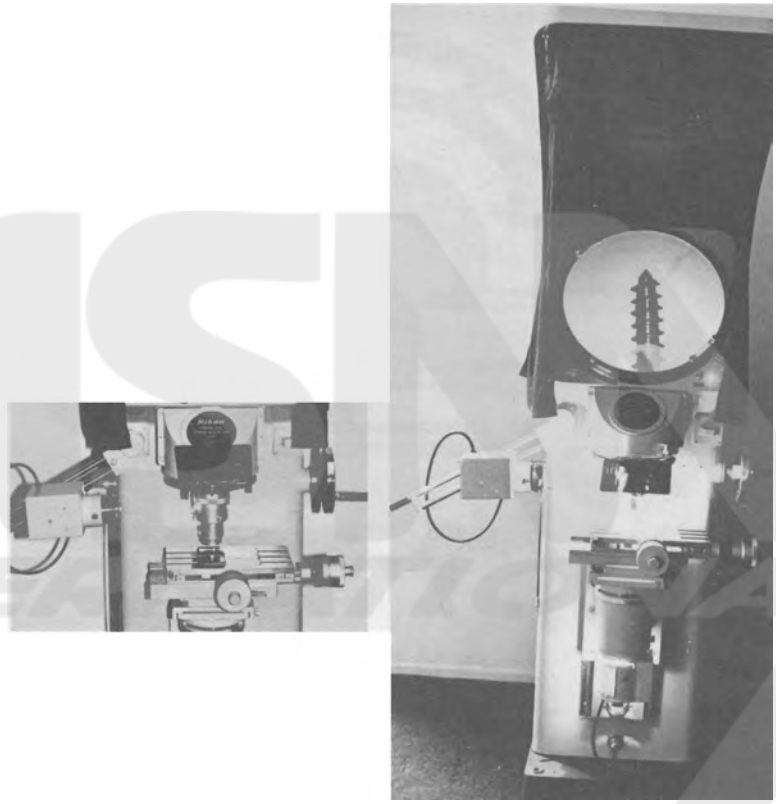


Fig. 13 Optical comparator. Source: Ref 1

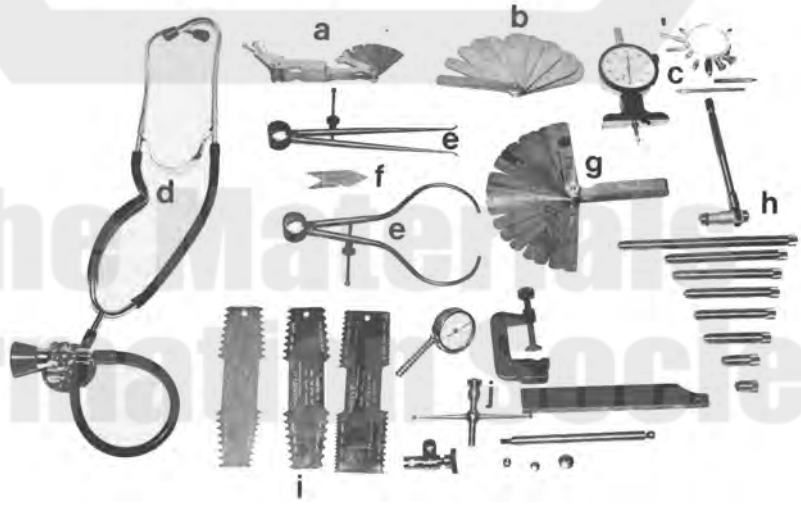


Fig. 14 Miscellaneous measuring devices: (a) gap measuring gage, (b) radius gage, (c) depth gage, (d) stethoscope, (e) inside and outside calipers, (f) center gage, (g) feeler gage, (h) inside micrometers, (i) thread profile gage, (j) dial indicator. Source: Ref 1

Miscellaneous Equipment

There are many other tools of the trade for visual examination. There are several other notable items that are definitely classified as visual examination items.

Stereoscopic Microscope. The stereoscopic microscope may well be the most important and widely used of all visual tools (Fig. 15). It allows three-dimensional viewing, clearly and sharply, to magnifications as high as 180×. There are several variations to consider with this equipment; they are lens combinations, the stand, and the zoom option. The lens combinations may be wholly or partly interchangeable. The ultimate use dictates the final choice, but for general-purpose work magnifications in the range of 5× to 50× are most popularly used.

The normal stand is similar to that for any upright microscope and is adequate if small parts are to be viewed. For most applications, however, the extension-arm type of a stand is much more versatile. The long extension arm allows the *scope* to swing out over fairly large parts to examine a specific area.

The zoom option is highly recommended. It allows a continuously variable range of magnification without changing lenses, by simply rotating a dial. For example, magnification may be varied from 5× to 50×. General observations can be made at 5×, and if something of interest requires more detail, a higher magnification, up to 50×, can rapidly be obtained.

Camera attachments are available with stereo-microscopes, but unless stereo pair photography is used the results are disappointing compared to the visual observation.

Mirrors. Another essential tool for the visual examiner is a mirror. Mirrors are available in all sizes and shapes, with and without lights. They are

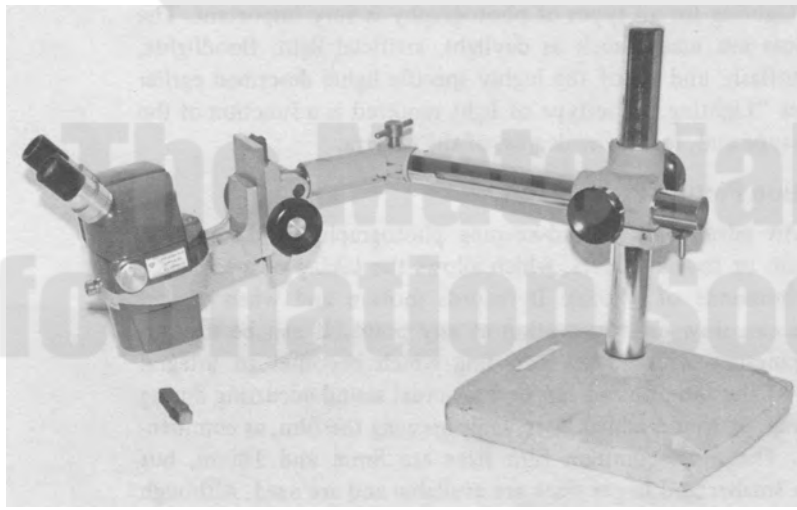
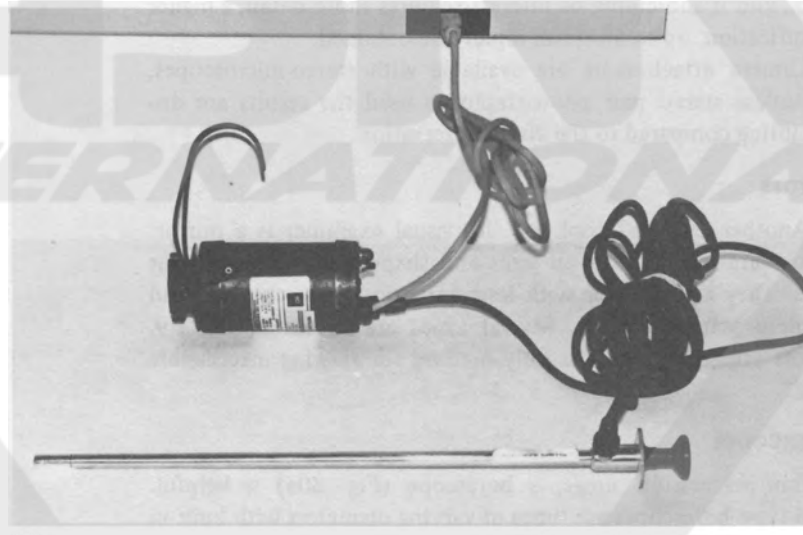


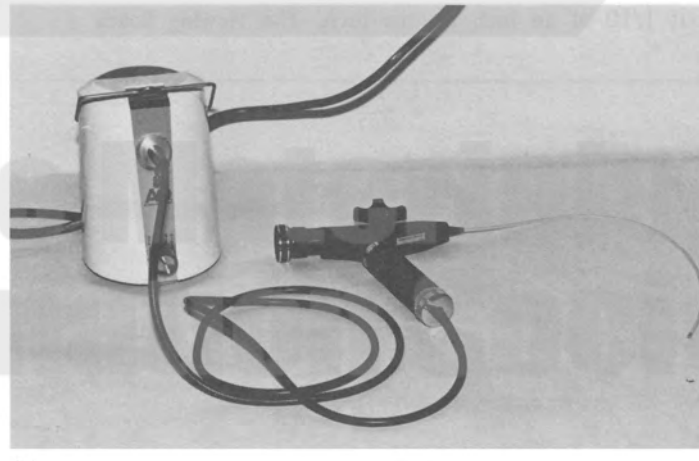
Fig. 15 Stereomicroscope with extension arm. Source: Ref 1

available with long extensions, swivel heads, and remotely actuated heads. Mirrors are sometimes the only method for viewing inaccessible areas.

Borescopes. For inaccessible areas, a borescope (Fig. 16a) is helpful. Rigid-type borescopes are tubes of varying diameters with built-in lens systems. They generally have built-in lighting systems. They come in fixed lengths as well as in the sectional form. Lengths vary from 150 mm (6 in.) to about 12 m (40 ft). Diameters will usually vary from about 2.5 to 25 mm (0.1 to 1 in.). The viewing heads offered are, for example, for straight-ahead viewing (0°), forward oblique viewing (30°), right-angle viewing (90°), retrograde viewing (110°), and panoramic viewing (180°).



(a)



(b)

Fig. 16 (a) Borescope, (b) fiber optic scope. Source: Ref 1

Eyepieces are available in monocular or binocular viewing. They come in both fixed focus and adjustable eyepiece focusing. They are also available with adaptation to video viewing on television screens and for photography. Both incandescent and halogen light sources can be utilized. Either plug-in or battery power sources can be used. Magnification can be varied by design, but is principally related to the distance between the subject and the objective head. These instruments are excellent for internal examination of long tubes, boreholes, internal combustion engine cylinders, castings, etc.

Fiber Optic Scopes. Another variety of the borescope is the fiber optic scope (Fig. 16b). These are very similar to borescopes, but have the ability to deform. The examining tube in this case is made up of thousands of carefully aligned glass fibers with an objective lens at one end and a magnifying eyepiece at the other. Since it is flexible, a fiber optic scope can *snake* its way around corners and along tortuous paths to examine inaccessible areas rigid scopes could not reach. These scopes are available up to 4.5 m (15 ft) long with a variety of accessories, including watertight viewing tubes.

Surface Finish Comparators. Many comparative test sheets are available to rate surface finishes (Fig. 17). The surface finish is often a requirement of visual examination. These comparator scales are available for rating machined surfaces, including turned, ground, lapped, milled, profiled, and electrical discharge machined (EDM). They are available for rating grit-blasted and sandblasted surfaces. Cast surfaces can also be rated.

Record-Keeping Methods

Verification of the results of visual examination requires some means of record keeping. Record keeping in its simplest form is accomplished by making written notation of the results. Since the making of written records is a somewhat slow and inconvenient process and may indeed end up being illegible, other methods are worthy of consideration. Several devices for record keeping are described below.

Voice Recorders. One simple and widely used device is the a portable voice recorder. With it, rapid note taking is possible, and the results can be transcribed later. The device may be electronic or a battery powered recorder using cassette tapes. The tapes can be used repeatedly or can be identified, replaced, and preserved. Although tape recorders are simple and accurate, they obviously cannot record sketches or visual depiction of results.

Photography is an excellent record-keeping method. It can be used very effectively in conjunction with written records or voice recordings. There have been many advances in this field in recent years, and volumes are available on the subject. The most commonly used devices for record-

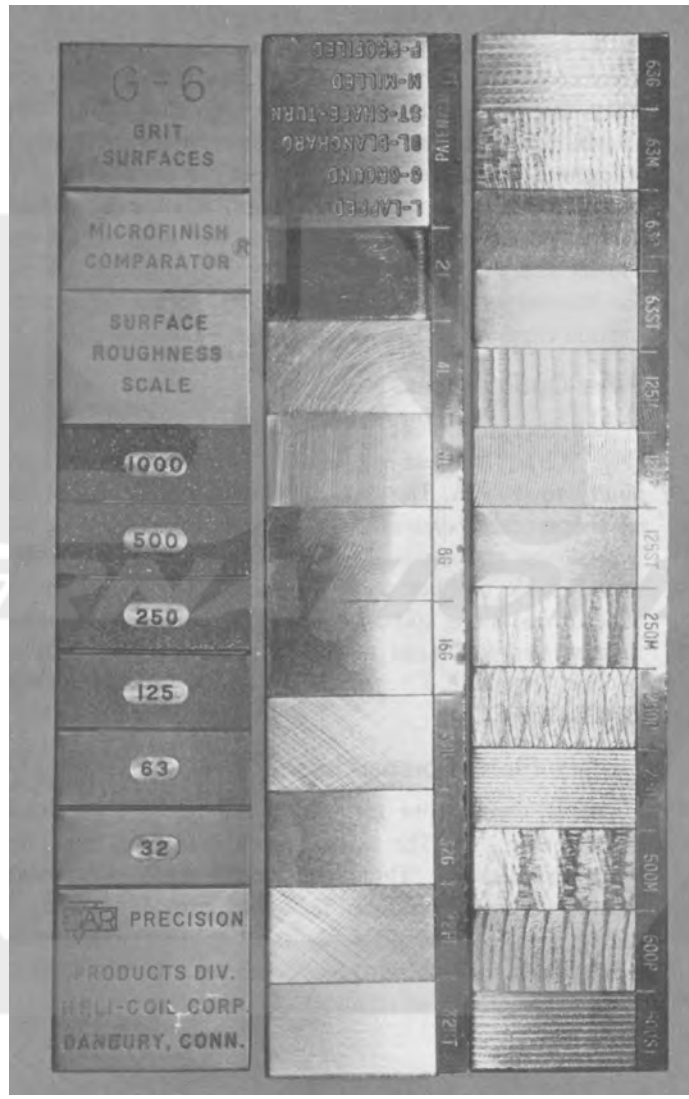


Fig. 17 Surface finish comparators. Source: Ref 1

ing the results of visual examination are the 35 mm single lens reflex camera, the digital cameras, the larger-format view cameras, and macro cameras. Because each of these devices has advantages and disadvantages, it is difficult to recommend only one for a general-purpose device. For this reason, each will be discussed in detail.

35 mm Film Cameras. Film cameras, although being displaced by digital cameras, may still be used. The 35 mm single lens reflex cameras may be the most widely used of the group. They are available with many accessories and lens types. Many have built-in light meters. They are very por-

table and produce excellent results. The operator looks directly through the lens in composing the picture, which prevents forgetting to remove the lens cap or having a thumb in the photo. At the time of exposure, a mirror device pivots out of the way, the exposure is made, and the viewing mirror returns to position. Lenses are interchangeable. Exposure times are variable and can be as little as $^{1}/_{1000}$ of a second. Apertures vary, allowing light-entrance control and variation in the depth of focus. Lighting can be sunlight, artificial, flash, or strobe-type high-intensity lighting. These cameras usually take 20- or 36- exposure films. Film is available in many types and speeds, in either black-and-white or color. Film is also available for prints or slide projection. Cost per picture is low due to the small size of the film. The principal disadvantage of this camera is the small size of the film, which requires enlarging. If enlarged too much, loss of detail and graininess can occur. Another disadvantage is that 20 or 36 exposures must be made before developing, to avoid wasting film. After all of the exposures, delay may be experienced while the film is developed and each frame separately enlarged and printed. But the speed, versatility, portability, and variety of high-quality equipment and film generally outweigh any disadvantages, making 35 mm cameras an excellent choice for a wide range of record-keeping chores.

Digital Cameras. Digital cameras have the advantage of producing rapid records. This ensures that the results are what is expected without the delay of shooting, developing, and printing a roll of film. The resolution in a digital image is determined by the number of pixels in the imaging sensor; the higher the number of pixels, the greater the image resolution. High resolution photographs are attainable with multiple megapixel sensor arrays. Image file sizes in the multiple megabyte range produce excellent photographs with enlargement capability.

View Cameras. View cameras are simply cameras with a larger format. Common sizes are 4 by 5 inches, 5 by 7 inches, and 8 by 10 inches. These cameras produce probably the best quality, most detailed photos. They are somewhat bulky and often require tripod support. Film is available in black and white and in color. Prints can be greatly enlarged with little loss of detail.

Macro cameras. are another form of the view camera. They are particularly adapted to produce magnified photos. This is accomplished by use of special lenses in conjunction with focusing bellows. These cameras are best used with smaller parts. Lighting for this equipment is important. At higher magnifications, the lens-to-subject distance is short and lighting is even more difficult. A variety of films are available for use with this equipment, including black-and-white cut film, color film, and all of the rapid film types.

Photography Lighting. Lighting for all types of photography is very important. The choices are many, such as daylight, artificial light, flood-

lights, photoflash, and all of the highly specific lights described earlier under the section "Lighting." The type of lighting required is a function of the film type and, to a lesser degree, of the camera.

Motion Picture Photography. An advance in record-keeping photography is the motion picture or movie camera, which allows the taking of a sequence of thousands of photos. It records motion and when played back can slow or stop motion at any point. It can be used in conjunction with sound recording which becomes an integral part of the film. Sound can be the actual sound occurring during filming, or sound added later while viewing the film, as commentary. Both black-and-white and color film are available.

Video Recorders. A record-keeping device that promises to be widely used in the future is the video tape recorder. The equipment has been used for many years by the television industry, but has only recently become available in more portable and inexpensive forms. The video tape recorder produces records similar to the movie camera. The principal difference is that the results can be viewed instantly without processing. They can be shown on any standard television or video screen, and the recording tapes can either be stored or erased and used again.

Macroetching

Various imperfections or defects invisible to the naked eye can often be detected by hot or cold acid etching. Since the cross-section usually provides more information than the longitudinal section, the general practice is to cut discs transversely; i.e., perpendicular to the hot working axis. To facilitate handling, disc thickness should generally be 25 mm (1 in.) or less. Longitudinal sectioning is used to study fibering, segregation, and inclusions. Common uses of macroetching are given below.

Solidification Structures. The structure resulting from solidification can be clearly revealed by macroetching. The macrostructure of a transverse disc cut from a small laboratory size steel ingot that was etched with 10% HNO_3 in water is shown in Fig. 18. At the mold surface, there is a small layer of very fine equiaxed grains. From this outer shell, large columnar grains grow inward toward the central, equiaxed region.

Billet and Bloom Macrostructures. The steelmaker uses hot acid etching on discs cut, with respect to the ingot location, from the top and bottom or the top, middle, and bottom of billets or blooms rolled from the first, middle, and last ingots teemed from the heat. If a disc reveals a rejectable condition, the billet material is rejected until the condition is removed.

Continuously Cast Steel Macrostructures. Continuous casting has become an important process for producing metals. Macroetching has been widely employed in the development of macroetching to evaluate the

Fig. 18 Cold etch of disc cut from small ingot (10% aqueous HNO_3). Source: Ref 2

influence of casting parameters on billet and slab quality and on the quality of the wrought product.

Forging Flow Lines. Macroetching is widely used to study metal flow patterns due to hot or cold working.

Grain or Cell Size. Macroetching usually reveals the as-cast grain structure, particularly when it is relatively coarse.

Alloy Segregation. Because most engineering alloys freeze over a range of temperatures and liquid compositions, the various elements in the alloy segregate during the solidification of ingots and castings. Segregation occurs over short distances, causing microsegregation, and over long distances, producing macrosegregation. *Microsegregation* is a natural result of dendritic solidification because the dendrites are purer in composition than the interdendritic matter. *Macrosegregation* manifests itself in a variety of forms–centerline segregation, negative cone of segregation, A- and V-type segregates, and banding. These phenomena are the result of the flow of solute enriched interdendritic liquid in the mushy zone during solidification which is a result of solidification shrinkage and gravitational forces.

Weldments. In any study of welds, the initial step invariably centers on the development of the weld macrostructure. The weld macrostructure is established by the type of process employed, the operating parameters, and the materials used. Thus, metallography is a key tool in weld quality studies. Key terms in describing the macrostructure of fusion welds are

the basic three components—the weld metal or *nugget*, the heat affected zone (HAZ), and the base metal.

Within the weld metal and the heat affected zone, there are changes in composition, grain size and orientation, microstructure, and hardness. Thus one observes significant variations in microstructure as the weldment is scanned.

Macroetching is frequently employed to determine the influence of various changes in weld parameters on the size and shape of the weld metal, on depth of penetration, on weld structure, and on hardness.

Response to Heat Treatment. Macroetching can also be used to determine the hardenability of various steel bars subjected to known heat treatment conditions. This procedure, coupled with hardness testing, was widely used prior to the adoption of hardenability analysis. As an illustration, Fig. 19 shows discs cut from round bars of AISI 1060 carbon steel ranging in size from a diameter of 20 to 65 mm (¾ to 2½ in.). The two smallest sizes were through hardened, that is, the center region contains more than 50% martensite, and the etch pattern was uniform. The other three sizes exhibit a case and core pattern, since the central region was unhardened. For this test, all bars were austenitized at 829 °C (1525°F), brine quenched, and then tempered at 149 °C (300°F). The bar length was twice the diameter, and the etched section was taken from the center.

Cold etching is also useful in studying the results of surface hardening treatments. Figure 20 shows the results of induction hardening of gear teeth made from AISI 1055 carbon steel. The areas hardened and the depth of the hardened zone are quite apparent.

Fig. 19 Macroetching (10% aqueous $HN0_3$) was used to reveal the extent of hardening in these AISI 1060 carbon steel round bars. Source: Ref 2

Fig. 20 The depth and extent of hardening in these induction hardened gear teeth made of AISI 1055 carbon steel was determined by macroetching with 10% aqueous HNO_3. Surface hardness was 53 to 54 HRC while the unhardened area was about 23 HRC. Source: Ref 2

ACKNOWLEDGMENT

This chapter was adapted from *Inspection of Metals, Volume 1, Visual Inspection* by R.C. Anderson, 1983, and *Metallography–Principles and Practice*, by G.F. Vander Voort, 1999.

REFERENCES

1. R.C. Anderson, *Inspection of Metals, Volume 1, Visual Inspection*, American Society for Metals, 1983
2. G.F. Vander Voort, *Metallography–Principles and Practice*, ASM International, 1999

CHAPTER 3

Coordinate Measuring Machines

THE COORDINATE MEASURING MACHINE (CMM) is used for three-dimensional inspection of both in-process and finished parts. Historically, CMMs have been largely used to measure and collect dimensional inspection data used to make acceptance or rejection decisions. Although CMMs continue to play this role, manufacturers are placing new emphasis on using CMMs to capture data from many sources and bring them together centrally where they can be used to control the manufacturing process more effectively and to prevent defective components from being produced. In addition, CMMs are also being used in entirely new applications; for example, reverse engineering and computer-aided design and manufacture (CAD/CAM) applications as well as innovative approaches to manufacturing, such as the flexible manufacturing systems, manufacturing cells, machining centers, and flexible transfer lines.

Important terminology for CMMs includes:

- *Ball bar:* A gage consisting of two highly spherical tooling balls of the same diameter connected by a rigid bar
- *Gage:* A mechanical artifact of high precision used either for checking a part or for checking the accuracy of a machine; a measuring device with a proportional range and some form of indicator, either analog or digital
- *Pitch:* The angular motion of a carriage, designed for linear motion, about an axis that is perpendicular to the motion direction and perpendicular to the yaw axis
- *Pixel:* The smallest element into which an image is divided, such as the dots on a television screen
- *Plane:* A surface of a part that is defined by three points

- *Repeatability:* A measure of the ability of an instrument to produce the same indication (or measured value) when sequentially sensing the same quantity under similar conditions
- *Roll:* The angular motion of a carriage, designed for linear motion, about the linear motion axis
- *Yaw:* The angular motion of a carriage, designed for linear motion, about a specified axis perpendicular to the motion direction. In the case of a carriage with horizontal motion, the specified axis should be vertical unless explicitly specified. For a carriage that does not have horizontal motion, the axis must be explicitly specified

CMM Operating Principles

A CMM is a multiaxial device with two to six axes of travel or reference axes, each of which provides a measurement output of position or displacement. Coordinate measuring machines are primarily characterized by their flexibility, being able to make many measurements without adding or changing tools. As products evolve, the same CMM can generally be used, depending on size and accuracy limitations, simply by altering software instead of altering equipment mechanics or electronics.

CMMs consist of the machine itself and its probes and moving arms for providing measurement input, a computer for making rapid calculations and comparisons (to blueprint specifications, for example) based on the measurement input, and the computer software that controls the entire system. In addition, the CMM has some means of providing output to the user (printer, plotter CRT) and/or to other machines in a complete manufacturing system. Coordinate measuring machines linked together in an overall inspection or manufacturing system are referred to as coordinate measuring systems.

The most important feature of the CMM is that it can rapidly and accurately measure objects of widely varying size and geometric configuration; for example, a particular part and the tooling for that part. Coordinate measuring machines can also readily measure the many different features of a part, such as holes, slots, studs, and weld-nuts, without needing other tools. Therefore, CMMs can replace the numerous hand tools used for measurement as well as the open plate and surface plate inspection tools and hard gages traditionally used for part measurement and inspection.

Coordinate measuring machines do not always achieve the rates of throughput or levels of accuracy possible with fixed automation type measuring systems. However, if any changes must be made in a fixed system for any reason; for example, a different measurement of the same part or measurement of a different part, making the change will be costly and time consuming. This is not the case with a CMM. Changes in the measurement or inspection routine of a CMM are made quickly and easily

by simply editing the computer program that controls the machine. The greater or more frequent the changes required, the greater the advantage of the CMM over traditional measuring devices. This flexibility, as well as the resulting versatility, is the principal advantage of the CMM.

CMM Measurement Techniques

A CMM takes measurements of an object within its work envelope by moving a sensing device called a probe along the various axes of travel until the probe contacts the object. The precise position of the contact is recorded and recorded as a measurement output of position or displacement (Fig. 1). The CMM is used to make numerous contacts, or hits, with the probe; using all axes of travel, until an adequate data base of the surfaces of the object has been constructed. Various features of an object require different quantities of hits to be accurately recorded. For example, a plane, surface, or circular hole can be recorded with a minimum of three hits.

Once repeated hits or readings have been made and stored, they can be used in a variety of ways through the computer and geometric measurement software. The data can be used to create a master program; for example, of the precise specifications for a part. They can also be compared (via the software) to stored part specification data or used to inspect production parts for compliance with specifications. A variety of other sophisticated applications are also possible using the same captured measurement data; for example, the reverse engineering of broken parts or the development of part specifications from handmade models.

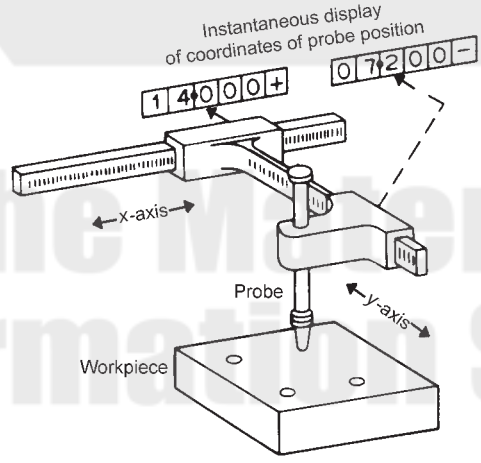


Fig. 1 Elements of a CMM showing typical digital position readout. The probe is positioned by brackets slid along two arms. Coordinate distances from one point to another are measured in effect by counting electronically the lines in gratings ruled along each arm. Any point in each direction can be set to zero, and the count is made in a plus or minus direction from there. Source: Ref 1

Coordinate Systems. The CMM registers the various measurements or hits it takes of an object by a system of coordinates used to calibrate the axes of travel. There are several coordinate systems in use. The most commonly used system is Cartesian, a three-dimensional, rectangular coordinate system; i.e., the same as that found on a machine tool. In this system, all axes of travel are square to one another. The system locates a position by assigning it values along the *x*, *y*, and *z* axes of travel.

Another system used is the polar coordinate system. This system locates a point in space by its distance or radius from a fixed origin point and the angle this radius makes with a fixed origin line. It is analogous to the coordinate system used on a radial arm saw or radial arm drill.

Types of Measurements

Fundamentally, CMMs measure the size and shape of an object and its contours by gathering raw data through sensors or probes. The data are then combined and organized through computer software programs to form a coherent mathematical representation of the object being measured, after which a variety of inspection reports can be generated. There are three general types of measurements for which CMMs are commonly used, geometric, contour, and specialized surface.

Geometric measurement deals with the elements commonly encountered every day–points, lines, planes, circles, cylinders, cones, and spheres. In practical terms, these two-dimensional and three-dimensional elements and their numerous combinations translate into the size and shape of various features of the part being inspected.

A CMM can combine the measurements of these various elements into a coherent view of the part and can evaluate the measurements. For example, it can gage the straightness of a line, the flatness of a plane surface, the degree of parallelism between two lines or two planes, the concentricity of a circle, the distance separating two features on a part, and so on. Geometric measurement clearly has broad application to many parts and to a variety of industries.

Contour measurement deals with artistic, irregular, or computed shapes, such as automobile fenders or aircraft wings. The measurements taken by a CMM can easily be plotted with an exaggerated display of deviation to simplify evaluation. Although contour measurements are generally not as detailed as geometric measurements, presenting as they do only the profile of an object with its vector deviation from the nominal or perfect shape, they too have broad application.

Specialized surface measurement deals with particular, recurring shapes, such as those found on gear teeth or turbine blades. In general, these shapes are highly complex, containing many contours and forms, and the part must be manufactured very precisely. Tight tolerances are absolutely critical. Because manufacturing accuracy is critical, measure-

ment is also highly critical, and a specialty in measuring these forms has evolved. By its nature, specialized surface measurement is applied to far fewer applications than the other two types.

CMM Capabilities

Coordinate measuring machines have the fundamental ability to collect a variety of different types of very precise measurements and to do so quickly, with high levels of repeatability and great flexibility. In addition, they offer other important capabilities based on computational functions.

Automatic Calculation of Measurement Data. The inclusion of a computer in the CMM allows the automatic calculation of such workpiece features as hole size, boss size, the distance between points, incremental distances, feature angles, and intersections. Prior to this stage of CMM development, an inspector had to write down the measurements he obtained and manually compare them to the blueprint. Not only is such a process subject to error, but it is relatively time consuming. While waiting for the results of the inspection, production decisions are delayed and parts (possibly not being produced to specifications) are being manufactured.

Compensation for Misaligned Parts. Coordinate measuring machines no longer require that the parts being measured be manually aligned to the coordinate system of the machine. The operator cannot casually place the part within the CMM work envelope. Once the location of the appropriate reference surface or line has been determined through a series of hits on the datum features of the part, the machine automatically references that position as its zero-zero starting point, creates an *x*, *y*, *z* part coordinate system, and makes all subsequent measurements relative to that point. In addition, the part does not have to be leveled within the work envelope. Just as the CMM will mathematically compensate if the part is rotationally misaligned, it will also compensate for any tilt in the part.

Multiple Frames of Reference. The CMM can also create and store multiple frames of reference or coordinate systems that allows features to be measured on all surfaces of an object quickly and efficiently. The CMM automatically switches to the appropriate new alignment system and zero point (origin) for each plane (face) of the part. The CMM can also provide axis and plane rotation automatically.

Probe Calibration. The CMM automatically calibrates for the size and location of the probe tip (contact element) being used. It also automatically calibrates each tip of a multiple-tip probe.

Part Program and Data Storage. The CMM stores the program for a given part so that the program and the machine are ready to perform whenever this part comes up for inspection. The CMM can also store the results of all prior inspections of a given part or parts so that a complete history of its production can be reconstructed. This same capability also provides the groundwork for statistical process control applications.

Part programs can also be easily edited, rather than completely rewritten, to account for design changes. When a dimension or a feature of a part is changed, only that portion of the program involving the workpiece revision must be edited to conform to part geometry.

Interface and Output. CMMs can be linked together in an overall system or can be integrated with other devices in a complete manufacturing system. The CMM can provide the operator with a series of prompts that explain what to do next and guide the operator through the complete measurement routine. Output is equally flexible. The user can choose the type and format of the report to be generated. Data can be displayed in a wide variety of charts and graphs. Inspection comments can be included in the hard copy report and/or stored in memory for analysis of production runs.

CMM Applications

Coordinate measuring machines are most frequently used in two major roles: quality control and process control. In the area of quality control, CMMs can generally perform traditional final part inspection more accurately, more rapidly, and with greater repeatability than traditional surface plate methods.

In process control, CMMs are providing new capabilities. Because of the on-line, real-time analytical capability of many CMM software packages, CMMs are increasingly used to monitor and identify evolving trends in production before scrap or out-of-spec parts are fabricated in the first place. Thus, the emphasis has shifted from inspecting parts and subsequently rejecting scrap parts at selected points along the production line to eliminating the manufacture of scrap parts altogether and producing in-tolerance parts 100% of the time.

In addition to these uses, there is a trend toward integrating CMMs into systems for more complete and precise control of production. Some shop hardened CMMs, also known as process control robots, are being increasingly used in sophisticated flexible manufacturing systems in the role of flexible gages.

Coordinate measuring machines can also be used as part of a CAD/CAM system. A CMM can measure a part, for example, and feed that information to the CAD/CAM program, which can then create an electronic model of the part. Going in the other direction, the model of the desired part in the CAD/CAM system can be used to create the part program automatically.

Types of CMMs

The ANSI/ASME B89 standard formally classifies CMMs into ten different types based on design. All ten types employ three axes of measure-

ment along mutually perpendicular guideways. They differ in the arrangement of the three movable components, the direction in which they move, and which one of them carries the probe, as well as where the workpiece is attached or mounted. However, among the many different designs of CMMs, each with its own strengths, weaknesses, and applications, there are only two fundamental types: vertical and horizontal. They are classified as such by the axis on which the probe is mounted and moves. The ANSI/ASME B89 Performance Standard classifies coordinate measuring machines as:

Vertical	Horizontal
Fixed-table cantilever	Moving ram, horizontal arm
Moving-table cantilever	Moving table, horizontal arm
Moving bridge	Fixed table, horizontal arm
Fixed bridge	
L-shaped bridge	
Column	
Gantry	

In addition, the two types of machines can be characterized to some degree by: the levels of accuracy they each achieve, although there is a considerable degree of overlap based on the design of an individual machine; the size of part they can handle; and, application. A general comparison of CMM types, applications, and levels of measurement accuracy a CMM user can expect is given in Table 1.

Vertical CMMs

Vertical CMMs, which have the probe or sensor mounted on the vertical *z*-axis, have the potential to be the most accurate type. Vertical CMMs in general can be more massive and can be built with fewer moving parts

Table 1 Typical CMM specifications

			Minimum measurement	
Application	**CMM Type**	**Bearing type**	**mm**	**in.**
Laboratory quality(a)				
Laboratory grade	Vertical, moving bridge	Air bearings	<0.003	<0.0001
Clean room	Vertical, moving bridge	Air bearings	<0.003	<0.0001
Production machines				
Open shop	Vertical, moving bridge	Air bearings	<0.013–0.025	<0.0005–0.001
Sheet metal	Horizontal (with fixed *x*–*y* axis)	Recirculating bearing packs	<0.025–0.050	<0.001–0.002
Clean room and shop	Vertical, moving bridge; all horizontal	Air/roller bearings	<0.050	<0.002

(a) These CMMs have specific environments into which they must be installed to maintain their rated accuracies. Source: Ref 1

than their horizontal counterparts. Therefore, they are more rigid and more stable. However, their limitation is the size of part they can conveniently handle, because the part to be measured must fit under the structural member from which the probe descends.

In an attempt to overcome this limitation, various designs have been produced. As a result, within the overall category of vertical CMM, there are designs utilizing moving or fixed bridges, cantilevers, gantries, and so on. However, each design approach is a compromise, because as the size of the work envelope of the machine and the travel distance along the axes increase, so also do the problems of maintaining rigidity and accuracy. A gantry design has proved to be an effective solution to the problem of increasing the capacity of a CMM while maintaining a high level of accuracy. A cantilever design, on the other hand, presents some inherent problems associated with isolating the CMM from floor vibrations and maintaining high precision due to the overhanging unsupported arm.

Cantilever type CMMs (Fig. 2) employ three movable components moving along mutually perpendicular guideways. The probe is attached to the first component, which moves vertically (*z*-direction) relative to the second. The second component moves horizontally (*y*-direction) relative to the third. The third component is supported at one end only, cantilever fashion, and moves horizontally (*x*-direction) relative to the machine base. The workpiece is supported on the worktable. A typical machine of this configuration is shown in Fig. 2(a). A modification of the fixed table can-

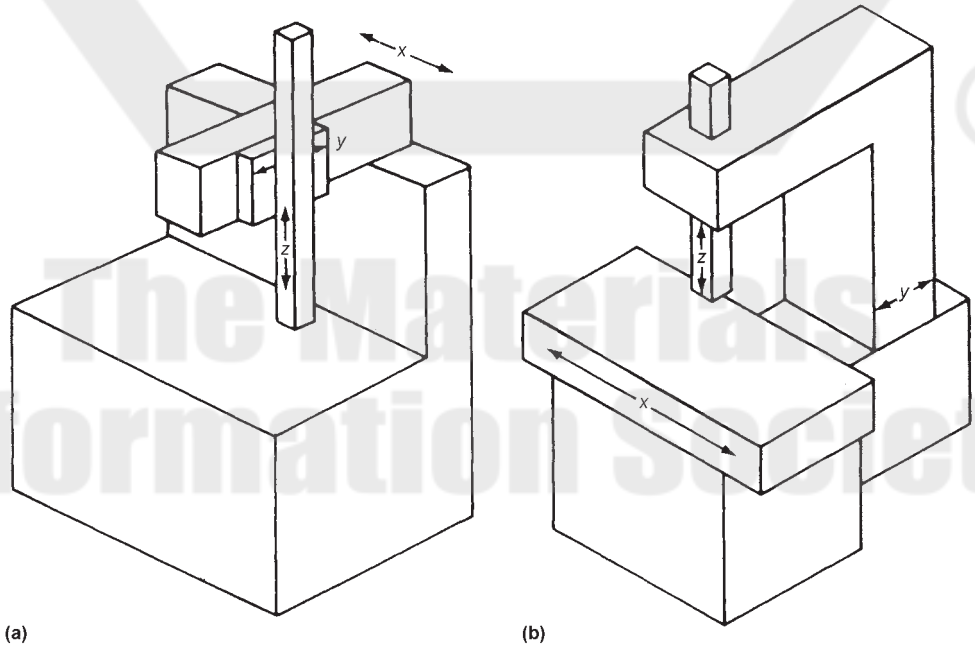


Fig. 2 Motion of components in cantilever type CMMs. (a) Fixed table. (b) Moving table. Source: Ref 1

tilever configuration is the moving table cantilever CMM shown in Fig. 2(b). The cantilever design provides openness and accessibility from three sides, making it popular for small, manual CMMs. Its cantilevered *y*-axis places a size limitation on this configuration. Because of the small *y*-*z* assembly, this configuration is lightweight and provides fast measuring speeds in direct computer control (DCC) applications.

Bridge type CMMs (Fig. 3) employ three movable components moving along mutually perpendicular guideways. The probe is attached to the first component, which moves vertically (*z*-direction) relative to the second. The second component moves horizontally (*y*-direction) relative to the third. The third component is supported on two legs that reach down to

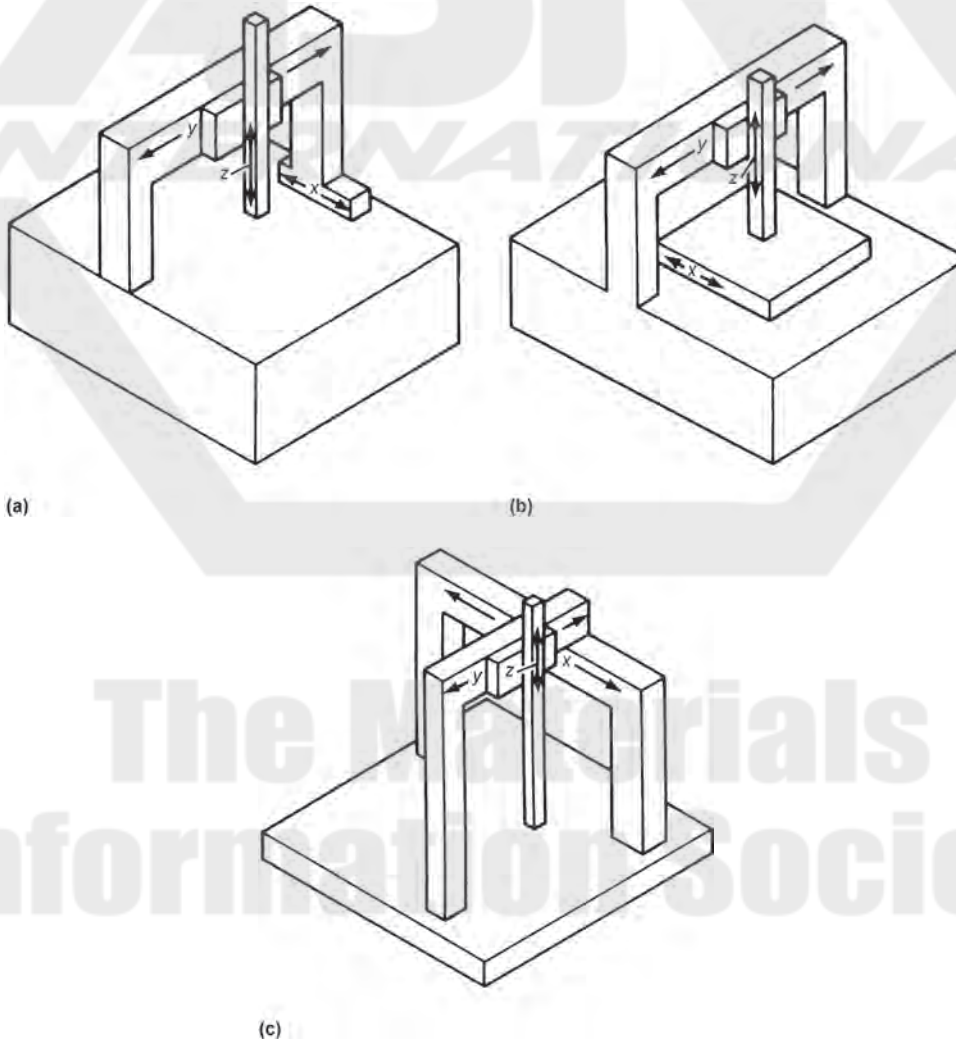


Fig. 3 Motion of components in bridge type CMMs. (a) Moving bridge. (b) Fixed bridge. (c) L-shaped bridge. Source: Ref 1

opposite sides of the machine base, and it moves horizontally (*x*-direction) relative to the base. The workpiece is supported on the base.

Moving Bridge. A typical moving bridge CMM is shown schematically in Fig. 3(a). This configuration accounts for 90% of all CMM sales. The moving bridge design overcomes the size limitations inherent in the cantilever design by providing a second leg, which allows for an extended *y*-axis. The second leg does, however, reduce access to the unit. The limitations of this configuration usually occur because of walking problems associated with retaining drive through just one leg. Higher speeds, achieved by increasing dynamic forces and reducing machine setting time, accentuate the problem. Vertical moving bridge CMMs can be controlled manually and with DCC hardware.

The fixed bridge configuration (Fig. 3b) provides a very rigid structure and allows a relatively light moving *x*-*z* structure that can achieve fast *x*-*z* moves. The moving table in larger machines can become massive, with decreased throughput capability. The influence of part weight on accuracy becomes a consideration for large parts.

L-Shaped Bridge. Another modification of the bridge configuration has two bridge-shaped components (Fig. 3c). One of these bridges is fixed at each end to the machine base. The other bridge, which is an inverted L-shape, moves horizontally (*x*-direction) on guideways in the fixed bridge and machine base.

The column CMM (Fig. 4) goes one step further than the fixed bridge in providing a very rigid *z*-axis configuration, and a two-axis saddle that allows movement in the horizontal (*x*-*y*) directions. High accuracy can be achieved with this design. However, as in the fixed bridge configuration, part mass and table considerations can restrict measuring volume and speed. Column CMMs are often referred to as universal measuring machines rather than CMMs by manufacturers. Column units are considered gage room instruments rather than production floor machines.

Gantry CMMs (Fig. 5) employ three movable components moving along mutually perpendicular guideways. The probe is attached to the probe quill, which moves vertically (*z*-direction) relative to a crossbeam. The probe quill is mounted in a carriage that moves horizontally (*y*-direction) along the crossbeam. The crossbeam is supported and moves in the *x*-direction along two elevated rails, which are supported by columns attached to the floor.

The gantry design has relatively restricted part access unless utilized in very large machines. The machine is physically large with respect to the size of the part. Large axis travels can be obtained, and heavy parts are not a problem, because the weight of the part can be decoupled from the measurement system by proper design of the machine base (foundation in a large machine). This is not as practical in smaller CMMs, and this configuration is most popular for large machines.

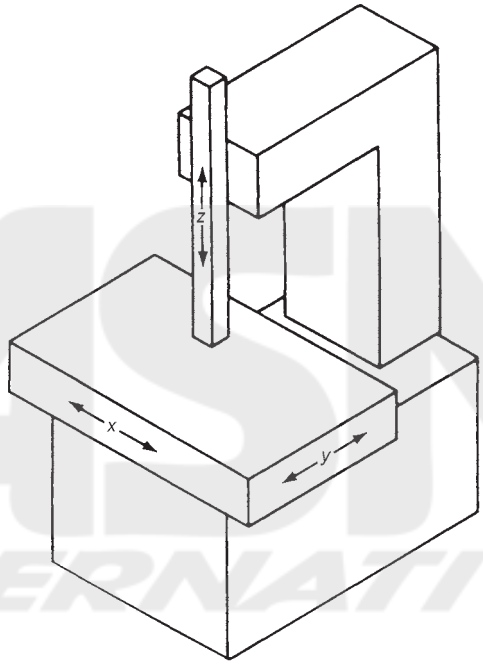


Fig. 4 Schematic of column CMM illustrating movement of probe, column, and table components. Source: Ref 1

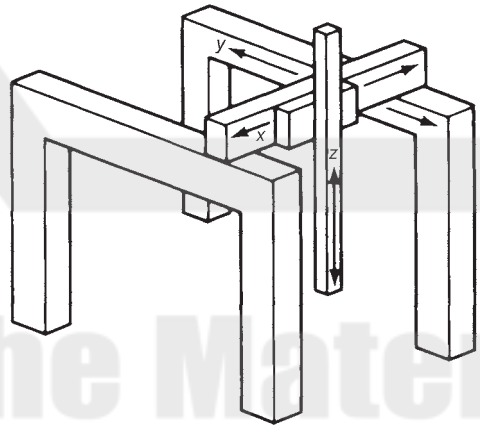


Fig. 5 Schematic of gantry CMM illustrating movement of probe, crossbeam, and elevated rails. Source: Ref 1

The gantry configuration was initially introduced in the early 1960s to inspect large parts, such as airplane fuselages, automobile bodies, ship propellers, and diesel engine blocks. The open design permits the operator to remain close to the part being inspected while minimizing the inertia of the moving machine parts and maintaining structural stiffness.

Horizontal CMMs

Horizontal CMMs (Fig. 6), which have the probe mounted on the horizontal *y*-axis, are generally used in applications in which large parts must be measured; for example, automobile bodies or airplane wings. Horizontal CMMs require no bridge over the part because the part is approached from the side. Therefore, there is substantially less restriction on the sizes of the parts that can be measured.

However, most horizontal CMMs do not measure to state-of-the-art levels of accuracy because of the high cost of achieving such accuracy in machines capable of handling large parts. In general, it is more cost effective to accept accuracy in the 0.050 mm (0.002 in.) range when gaining

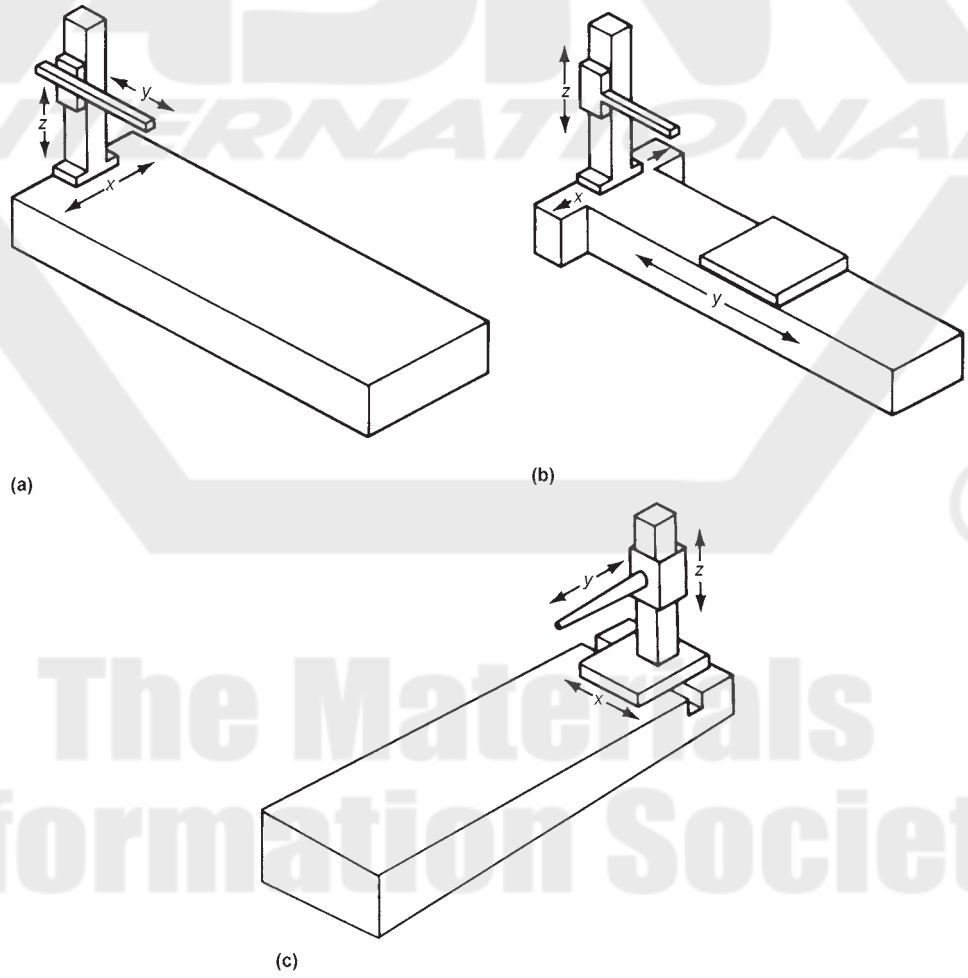


Fig. 6 Schematic illustrating three types of horizontal arm CMMs. (a) Moving ram. (b) Moving table. (c) Fixed table. Length of table/base is usually two to three times the width. Units with bases capable of accommodating two to three interstate buses situated end-to-end have been built for major automotive manufacturers. Source: Ref 1

the part handling capabilities of large horizontal machines and to use vertical designs when finer measurement is demanded.

In any CMM design, fewer moving parts and joints will result in higher potential levels of accuracy. This principle has been applied to a class of horizontal CMM called process control robots. In these units, fixed members that move together provide the flexibility and capacity of a horizontal CMM, along with the higher accuracy of a vertical design. The horizontal direction of attack makes these CMMs the logical design for production applications in which horizontal machine tools are used.

In choosing a CMM, the buyer must take into consideration not only the degree of accuracy required but also the location of the unit and the measurements to be taken. In general, when dealing with smaller parts where measurements of very high accuracy are required, the potential user is likely to be best served by a machine located in a clean room environment. When on-line process control is desired, the appropriate shop hardened CMM that can be located in the shop itself is the best solution.

Horizontal Arm CMMs. Several different types of horizontal arm CMMs are available. As with all CMMs, the horizontal arm configuration employs three movable components moving along mutually perpendicular guideways.

Horizontal arm CMMs are used to inspect the dimensional and geometric accuracy of a broad spectrum of machined or fabricated workpieces. Utilizing an electronic probe, these machines check parts in a mode similar to the way they are machined on horizontal machine tools. They are especially suited for measuring large gear cases and engine blocks, for which high precision bore alignment and geometry measurements are required. Four-axis capability can be obtained by incorporating a rotary table.

Horizontal arms for large machines have a lower profile than vertical arms. For some applications, horizontal access is desirable. For others, it is restrictive and a rotary table is usually required, thus increasing the cost.

Moving Ram Type. In this design, the probe is attached to the horizontal arm, which moves in a horizontal *y*-direction (Fig. 6a). The ram is encased in a carriage that moves in a vertical (*z*) direction and is supported on a column that moves horizontally (*x*-direction) relative to the base.

Moving Table Type. In this configuration, the probe is attached to the horizontal arm, which is permanently attached at one end only to a carriage that moves in a vertical (*z*) direction (Fig. 6b) on the column. The arm support and table move horizontally (*x*- and *y*-directions) relative to the machine base. The moving table horizontal arm CMM unit is even more versatile because of the introduction of a rotary moving table.

Fixed Table Type. In the fixed table version, the probe is attached to the horizontal arm, which is supported cantilever style at the arm support and moves in a vertical (*z*) direction (Fig. 6c). The arm support moves hori-

zontally (*x*- and *y*-directions) relative to the machine base. Parts to be inspected are mounted on the machine base.

ACKNOWLEDGMENT

This chapter was adapted from Coordinate Measuring Machines by D.H. Genest in *Nondestructive Evaluation and Quality Control*, Volume 17, *ASM Handbook* 1992.

REFERENCES

1. D.H. Genest, Coordinate Measuring Machines, *Nondestructive Evaluation and Quality Control*, Vol 17, *ASM Handbook*, ASM International, 1992, p 20

CHAPTER 4

Machine Vision

MACHINE VISION emerged as an important new technique for industrial inspection and quality control in the early 1980s. When properly applied, machine vision can provide accurate and inexpensive 100% inspection of workpieces, dramatically increasing product quality. Machine vision is also used as an in-process gaging tool for controlling the process and correcting trends that could lead to the production of defective parts. Consequently, manufacturers in a variety of industries have investigated this important technology, regardless of the products being manufactured. The automotive and electronics industries are the largest users of machine vision systems.

Machine vision, sometimes referred to as computer vision or intelligent vision, is a means of simulating the image recognition and analysis capabilities of the human eye/brain system with electronic and electromechanical techniques. A machine vision system senses information about an image and analyzes the information to make a useful decision about its content; in much the same way, the eye acts as the body's image sensor, with the brain analyzing this information and taking action based on the analysis.

Therefore, a machine vision system includes both visual sensing and interpretive capabilities. An image sensing device, such as a vidicon camera or a charge-coupled device (CCD) image sensor, is nothing more than a visual sensor that receives light through its lens and converts this light into electrical signals. When a data processing device, such as a microcomputer, is used, these electrical signals can be refined and analyzed to provide an interpretation of the scene that generated the signals. This information can then be used as a basis for taking an appropriate course of action. The entire process of image formation, analysis, and decision making is referred to as machine vision.

This ability to acquire an image, analyze it, and then make an appropriate decision is extremely useful in inspection and quality control applica-

tions. It enables machine vision to be used for a variety of functions, including:

- Identification of shapes
- Measurement of distances and ranges
- Gaging of sizes and dimensions
- Determining orientation of parts
- Quantifying motion
- Detecting surface shading

These functional capabilities allow users to employ machine vision systems for cost-effective and reliable 100% inspection of workpieces.

The analogy of human eye/brain system is helpful in understanding machine vision, but the human eye/brain system is extremely complex and operates in ways and at data rates much different from those of commercial machine vision systems. Humans are more flexible and often faster than machine vision systems. On the other hand, machine vision systems provide capabilities not achievable by humans, particularly with respect to consistency and reliability. Table 1 compares human and machine vision capabilities, and Table 2 evaluates the performance of each. In addition,

Table 1 Comparison of machine and human vision capabilities

Capabilities	Machine vision	Human vision
Distance	Limited capabilities	Good qualitative capabilities
Orientation	Good for two dimensions	Good qualitative capabilities
Motion	Limited; sensitive to image blurring	Good qualitative capabilities
Edges/regions	High-contrast image required	Highly developed
Image shapes	Good quantitative measurements	Qualitative only
Image organization	Special software needed; limited capability	Highly developed
Surface shading	Limited capability with gray scale	Highly developed
Two-dimensional interpretation	Excellent for well-defined features	Highly developed
Three-dimensional interpretation	Very limited capabilities	Highly developed
Overall	**Best for qualitative measurement of structured scene**	**Best for qualitative interpretation of complex, unstructured scene**

Source: Ref 1

Table 2 Evaluation of machine and human vision capabilities

Performance criteria	Machine vision	Human vision
Resolution	Limited by pixel array size	High resolution capability
Processing speed	Fraction of a second per image	Real-time processing
Discrimination	Limited to high-contrast images	Very sensitive discrimination
Accuracy	Accurate for part discrimination based on quantitative differences; accuracy remains consistent at high production volume.	Accurate at distinguishing qualitative differences; may decrease at high volume
Operating cost	High for low volume; lower than human vision at high volume	Lower than machine at low volume
Overall	**Best at high production volume**	**Best at low or moderate production volume**

Source: Ref 1

machine vision systems can detect wavelengths in the ultraviolet and infrared ranges, while the human eye is limited to wavelengths in the visible range (Fig. 1).

As indicated in Tables 1 and 2, neither machine vision nor human vision is clearly superior in all applications. Human vision is better for the low speed, qualitative interpretation of complex, unstructured scenes. An example in which human vision is superior to machine vision is the inspection of automobile body surfaces for paint quality. Human vision can easily and quickly detect major flaws, such as paint sagging, scratches, or unpainted areas, while this same task would be much more difficult and time consuming with machine vision techniques.

On the other hand, machine vision is better suited to the high speed measurement of quantitative attributes in a structured environment. Thus, machine vision is very good at inspecting the masks used in the production of microelectronic devices and at measuring basic dimensions for machined workpieces. Machine vision can not only perform these types of inspection better than humans but can do so reliably, without the fatigue and the errors that confront humans doing these types of repeated inspection tasks.

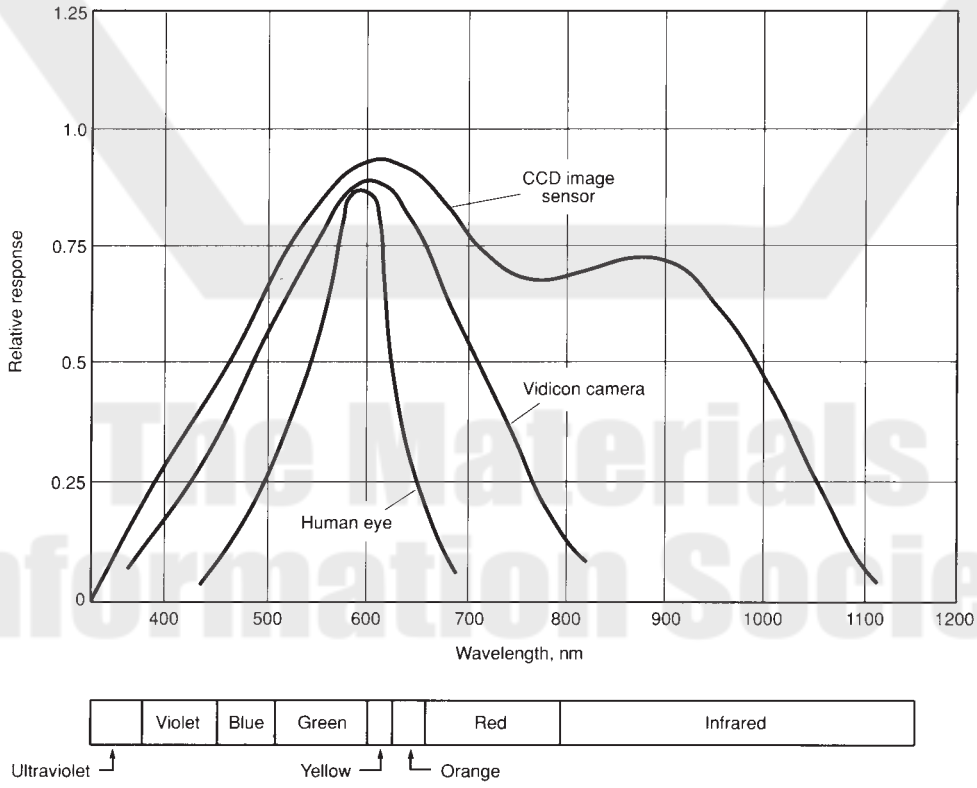


Fig. 1 Spectral response of the human eye, vidicon camera, and CCD image sensor. Source: Ref 2

Machine vision also has several additional important characteristics. First, it is a noncontact measurement technique. This is particularly important if the workpiece is fragile or distorts during contact, or if the workpiece would be contaminated or damaged if it were touched. Second, machine vision can be very accurate. Although accuracy is a function of many variables, including camera resolution, lens quality, field of view, and workpiece size, machine vision systems are often used to make measurements with an accuracy of ±3 μm (±120 μin.) or better. Third, machine vision can perform these functions at relatively large standoff distances–up to 1 m (3 ft) or more in some applications. Finally, these capabilities can be provided at relatively low cost. The price of a machine vision system may range from $5000 to $500,000, depending on the specific application and the capabilities of the system, but the typical price is less than $50,000. Collectively, these characteristics of machine vision provide the user with a capability that, for many applications, cannot be matched by human vision or other sensor or inspection technologies.

Machine Vision Process

To understand the capabilities and limitations of machine vision, it is useful to examine how a machine vision system operates. The key components of a machine vision system are shown in Fig. 2, and the process is illustrated in Fig. 3.

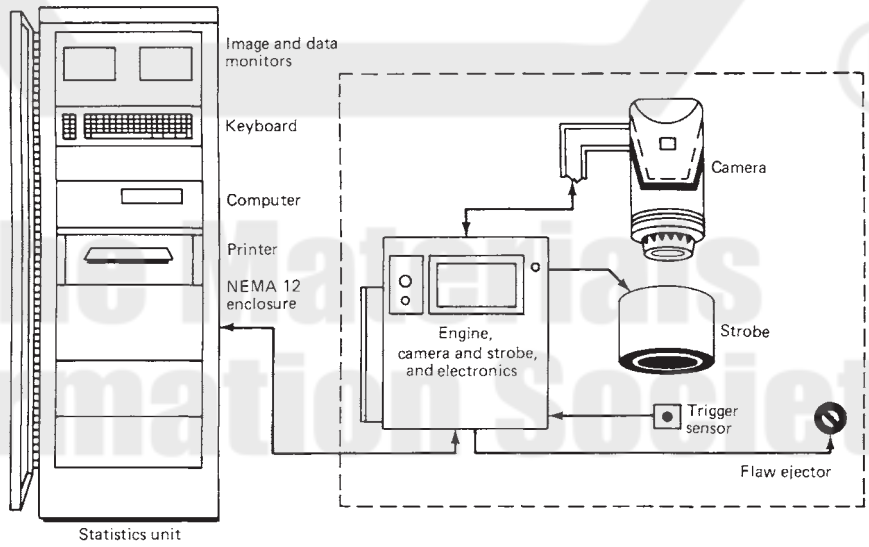


Fig. 2 Schematic illustrating the key components of a machine vision system. Source: Ref 3

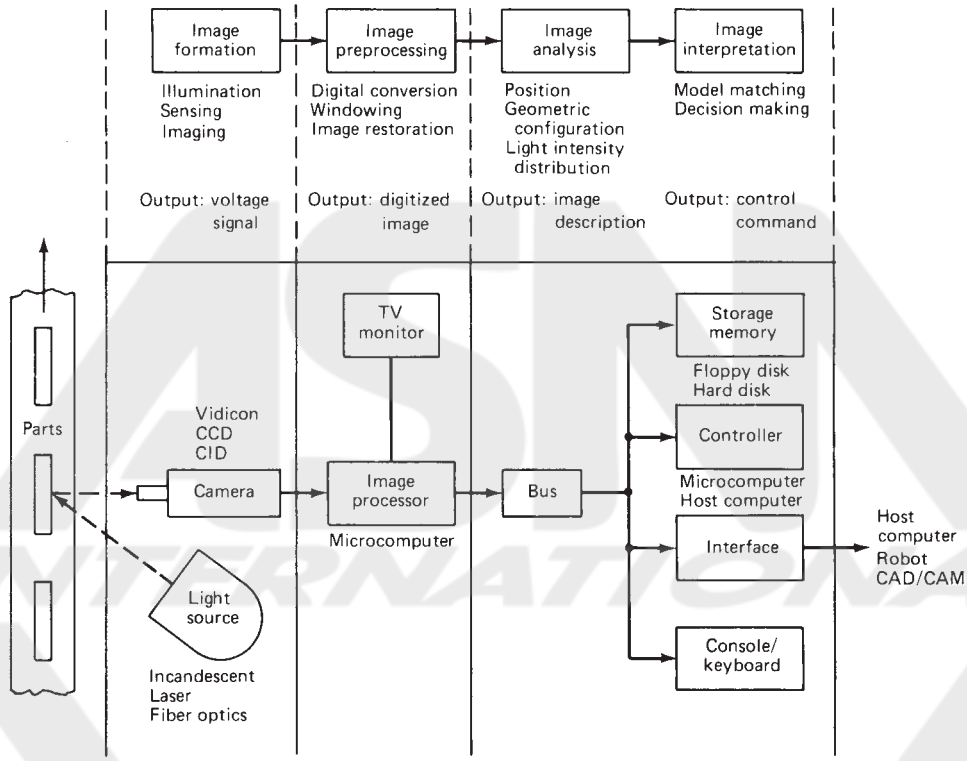


Fig. 3 Overview of the machine vision process. CID, charge injected device. Source: Ref 1

The machine vision process consists of four basic steps:

1. An image of the scene is formed—image formation
2. The image is processed to prepare it in a form suitable for computer analysis—image preprocessing
3. The characteristics of the image are defined and analyzed—image analysis
4. The image is interpreted, conclusions are drawn, and a decision is made or action taken, such as accepting or rejecting a workpiece—image interpretation

Image Formation

The first step of the machine vision process begins with the formation of an image, typically of a workpiece being inspected or operated on. Image formation is accomplished by using an appropriate sensor, such as a vidicon camera, to collect information about the light being generated by the workpiece.

The light being generated by the surface of a workpiece is determined by a number of factors, including the orientation of the workpiece, its surface finish, and the type and location of lighting being employed. Typical light sources include incandescent lights, fluorescent tubes, fiber optic bundles, arc lamps, and strobe lights. Laser beams are also used in some special applications, such as triangulation systems for measuring distances. Polarized or ultraviolet light can also be used to reduce glare or to increase contrast.

Proper Illumination. Correct placement of the light source is extremely important because it has a major effect on the contrast of the image. Several commonly used illumination techniques are illustrated in Fig. 4. When a simple silhouette image is all that is required, backlighting of the workpiece can be used for maximum image contrast. If certain key features on the surface of the workpiece must be inspected, front lighting would be used. If a three-dimensional feature is inspected, side lighting or structured lighting may be required. In addition to proper illumination, fixturing of the workpiece may also be required to orient the part properly and to simplify the rest of the machine vision process.

Once the workpiece or scene has been properly arranged and illuminated, an image sensor is used to generate the electronic signal representing the image. The image sensor collects light from the scene (typically through a lens) and then converts the light into electrical energy by using a photosensitive target. The output is an electrical signal corresponding to the input light.

Most image sensors used in industrial machine vision systems produce signals representing two-dimensional arrays or scans of the entire image, such as those formed by conventional television cameras. Some image sensors, however, generate signals using one-dimensional or linear arrays that must be scanned numerous times in order to view the entire scene.

Vidicon Camera. The most common image sensor in the early machine vision systems was the vidicon camera, which was extensively used in closed circuit television systems and consumer video recorders. An image is formed by focusing the incoming light through a series of lenses onto the photoconductive faceplate of the vidicon tube. An electron beam within the tube scans the photoconductive surface and produces an analog output voltage proportional to the variations in light intensity for each scan line of the original scene. Normally, the output signal conforms to commercial television standards–525 scan lines interlaced into 2 fields of 262.5 lines and repeated 30 times per second.

The vidicon camera has the advantage of providing a great deal of information about a scene at very fast speeds and at relatively low cost. However, vidicon cameras do have several disadvantages. They tend to distort the image due to their construction and are subject to image burn-in on the photoconductive surface. Vidicon cameras also have limited service lives and are susceptible to damage from shock and vibration.

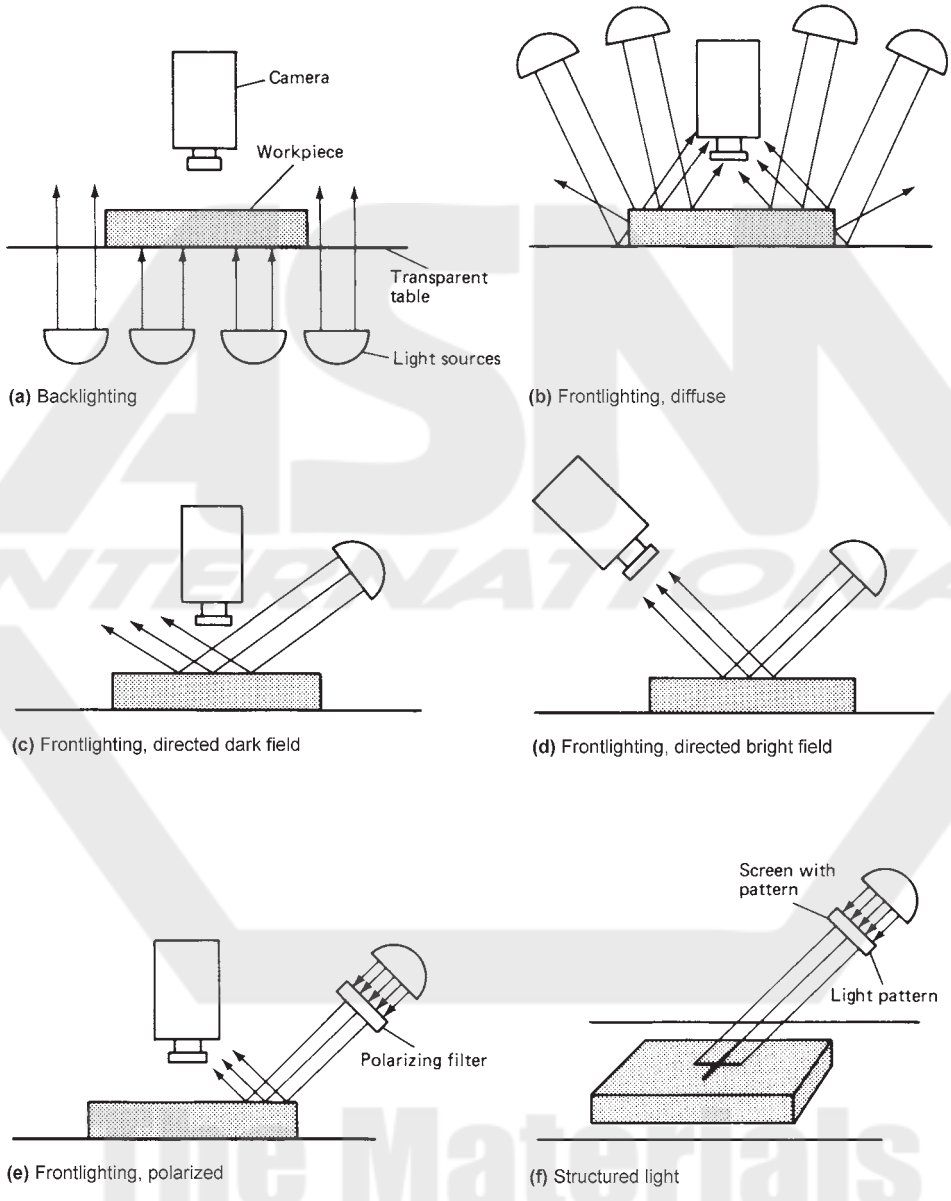


Fig. 4 Schematics of commonly used illumination techniques for machine vision systems. Source: Ref 1

Solid State Cameras. Most state-of-the-art machine vision systems use solid-state cameras, which employ charge-coupled (Fig. 5) or charge-injected device image sensors. These sensors are fabricated on silicon chips using integrated circuit technology. They contain matrix or linear arrays of small, accurately spaced photosensitive elements. When light passing through the camera lens strikes the array, each detector converts the light falling on it into a corresponding analog electrical signal. The

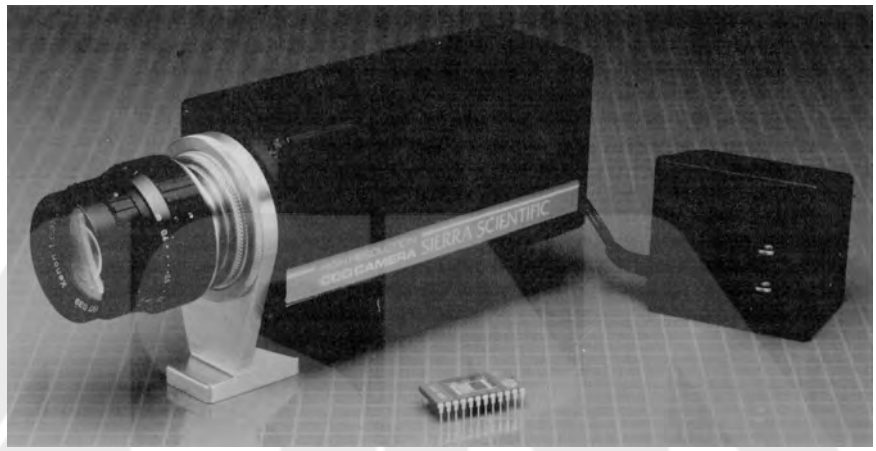


Fig. 5 Typical CCD image sensor. Courtesy of Sierra Scientific Corporation. Source: Ref 3

entire image is thus broken down into an array of individual picture elements known as pixels. The magnitude of the analog voltage for each pixel is directly proportional to the intensity of light in that portion of the image. This voltage represents an average of the light intensity variation of the area of the individual pixel. Charge-coupled and charge-injected device arrays differ primarily in how the voltages are extracted from the sensors.

Typical matrix array solid state cameras have 256 × 256 detector elements per array, although a number of other configurations are also popular. The output from these solid state matrix array cameras may or may not be compatible with commercial television standards. Linear array cameras typically have 256 to 1024 or more elements. The use of a linear array necessitates some type of mechanical scanning device (such as a rotating mirror) or workpiece motion (such as a workpiece traveling on a conveyor) to generate a two-dimensional representation of an image.

Selection of a solid-state camera for a particular application will depend on a number of factors, including the resolution required, the lenses employed, and the constraints imposed by lighting cost, and so on. Solid-state cameras offer several important advantages over vidicon cameras. In addition to being smaller than vidicon cameras, solid-state cameras are also more rugged. The photosensitive surfaces in solid-state sensors do not wear out with use as they do in vidicon cameras. Because of the accurate placement of the photo detectors, solid-state cameras also exhibit less image distortion. On the other hand, solid-state cameras are usually more expensive than vidicon cameras, but this cost difference is narrowing.

Although most industrial machine vision systems use image sensors of the types described above, some systems use special purpose sensors for

unique applications. This would include, for example, specialty sensors for weld seam tracking and other sensor types, such as ultrasonic sensors.

Image Preprocessing

The initial sensing operation performed by the camera results in a series of voltage levels that represent light intensities over the area of the image. In the second step of the machine vision process, this preliminary image must then be processed so that it is presented to the microcomputer in a format suitable for analysis. A camera typically forms an image 30 to 60 times per second, or once every 33 to 17 ms. At each time interval, the image is captured, or frozen, for processing by an image processor. The image processor, which is typically a microcomputer, transforms the analog voltage values for the image into corresponding digital values by means of an analog-to-digital converter. The result is an array of digital numbers that represent a light intensity distribution over the image area. This digital pixel array is then stored in memory until it is analyzed and interpreted.

Depending on the number of possible digital values that can be assigned to each pixel, vision systems can be classified:

- *Binary System.* The voltage level for each pixel is assigned a digital value of 0 or 1, depending on whether the magnitude of the signal is less than or greater than some predetermined threshold level. The light intensity for each pixel is considered to be either white or black, depending on how light or dark the image is.
- *Gray Scale System.* Like the binary system, the gray scale vision system assigns digital values to pixels, depending on whether or not certain voltage levels are exceeded. The difference is that a binary system allows two possible values to be assigned, while a gray scale system typically allows up to 256 different values. In addition to white or black, many different shades of gray can be distinguished. This greatly increased refinement capability enables gray scale systems to compare objects on the basis of such surface characteristics as texture, color, or surface orientation, all of which produce subtle variations in light intensity distributions. Gray scale systems are less sensitive to the placement of illumination than binary systems, in which threshold values can be affected by lighting.

Most commercial vision systems are binary. For simple inspection tasks, silhouette images are adequate; for example, to determine if a part is missing or broken. However, gray scale systems are used in many applications that require a higher degree of image refinement. The effects of gray scale digitization on the surface of an integrated circuit module are shown in Fig. 6.

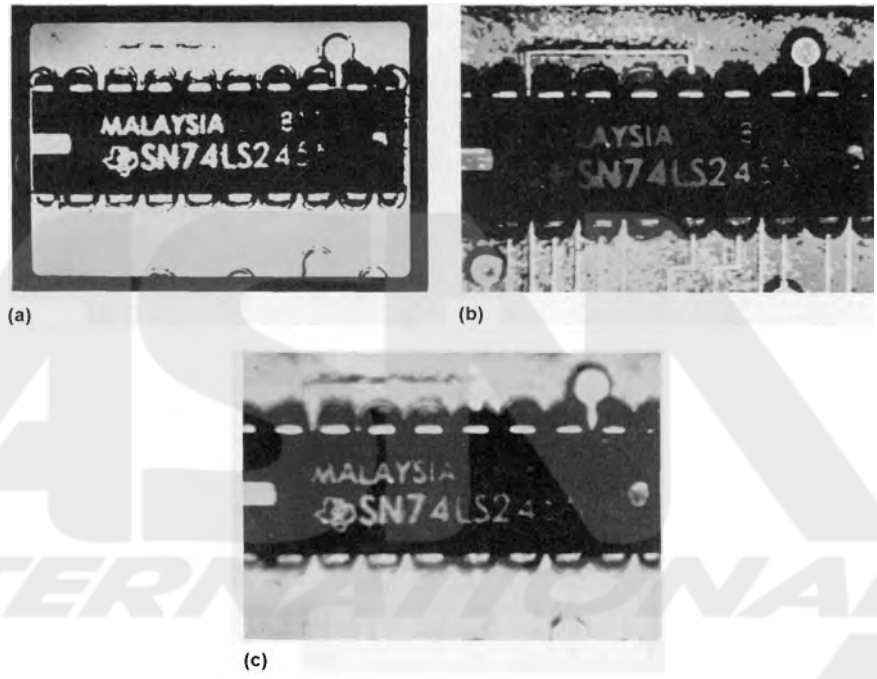


Fig. 6 Gray scale digitization of an IC module on a printed circuit board. (a) Binary. (b) 8-level. (c) 64-level. Courtesy of Cognex Corporation. Source: Ref 3

One of the most fundamental challenges to the widespread use of true gray scale systems is the greatly increased computer processing requirements relative to those of binary systems. A 256 × 256 pixel image array with up to 256 different values per pixel will require over 65,000 8-bit storage locations for analysis. At a speed of 30 images per second, the data processing requirement becomes very large, which means the time required to process large amounts of data can be significant. Ideally, a vision system should be capable of the real-time processing and interpretation of an image, particularly when the system is used for on-line inspection or the guidance and control of equipment such as robots.

One approach to reducing the amount of data to be processed; therefore, substantially reducing the time, is a technique known as windowing. This process creates an electronic mask around a small area of an image to be studied. Only the pixels that are not blocked out will be analyzed by the computer. This technique is especially useful for such simple inspection applications as determining whether or not a certain part has been attached to another part. Rather than process the entire image, a window can be created over the area where the attached part is expected to be located. By simply counting the number of pixels of certain intensity within the window, a quick determination can be made as to whether or not the part is

present. A window can be virtually any size, from one pixel up to a major portion of the image.

Another way in which the image can be prepared in a more suitable form during the preprocessing step is through the techniques of image restoration. Very often an image suffers various forms of degradation, such as blurring of lines or boundaries, poor contrast between image regions, or the presence of background noise. There are several possible causes of image degradation, including:

- Motion of the camera or object during image formation
- Poor illumination or improper placement of illumination
- Variations in sensor response
- Defects or poor contrast on the surface of the subject, such as deformed letters on labels or overlapping parts with similar light intensities

Techniques for improving the quality of an image include:

- *Constant brightness addition* involves simply adding a constant amount of brightness to each pixel. This improves the contrast in the image.
- *Contrast stretching* increases the relative contrast between high and low intensity elements by making light pixels lighter and dark pixels darker.
- *Fourier-domain processing* is a powerful technique based on the principle that changes in brightness in an image can be represented as a series of sine and cosine waves. These waves can be described by specifying amplitudes and frequencies in a series of equations. By breaking the image down into its sinusoidal components, each component image wave can be acted upon separately. Changing the magnitude of certain component waves will produce a sharper image that results in a less blurred image, better defined edges or lines, greater contrast between regions, or reduced background noise.

Some machine vision systems perform additional operations as part of the preprocessing function to facilitate image analysis or to reduce memory storage requirements. These operations, which significantly affect the design, performance, and cost of vision systems, differ according to the specific system and are largely dependent on the analysis technique employed in later stages of the process. Two commonly used preprocessing operations are:

- *Edge Detection.* An edge is a boundary within an image where there is a dramatic change in light intensity between adjacent pixels. These boundaries usually correspond to the real edges on the workpiece being examined by the vision system and are very important for applications such as the inspection of part dimensions. Edges are usually

determined by using one of a number of different gradient operators that mathematically calculate the presence of an edge point by weighting the intensity value of pixels surrounding the point. The resulting edges represent a skeleton of the outline of the parts contained in the original image.

Some vision systems include thinning, gap filling, and curve smoothing to ensure that the detected edges are only one pixel wide, continuous, and appropriately shaped. Rather than storing the entire image in memory, the vision system stores only the edges or some symbolic representation of the edges, thus dramatically reducing the amount of memory required.

- *Run length encoding* is another preprocessing operation used in some vision systems. This operation is similar to edge detection in binary images. In run length encoding, each line of the image is scanned, and transition points from black-to-white or white-to-black are noted, along with the number of pixels between transitions. These data are then stored in memory instead of the original image, and serve as the starting point for the image analysis. One of the earliest and most widely used vision techniques, originally developed by Stanford Research Institute and known as the SRI algorithms, uses run length encoded imaged data.

Image Analysis

The third general step in the vision sensing process consists of analyzing the digital image that has been formed so that conclusions can be drawn and decisions made. This is normally performed in the central processing unit of the system. The image is analyzed by describing and measuring the properties of several image features. These features may belong to the image as a whole or to regions of the image. In general, machine vision systems begin the process of image interpretation by analyzing the simplest features and then adding more complicated features until the image is clearly identified. A large number of different techniques are either used for use in commercial vision systems to analyze the image features describing the position of the object, its geometric configuration, and the distribution of light intensity over its visible surface.

Determining the position of a part with a known orientation and distance from the camera is one of the simpler tasks a machine vision system can perform. For example, consider the case of locating a round washer lying on a table so that it can be grasped by a robot. A stationary camera is used to obtain an image of the washer. The position of the washer is then determined by the vision system through an analysis of the pattern of the black and white pixels in the image. This position information is transmitted to the robot controller, which calculates an appropriate trajectory for the robot arm. However, in many cases, neither the distance between the

part and the camera nor the part orientation is known, making the task of the machine vision system much more difficult.

Object-Camera Distance Determination. The distance, or range, of an object from a vision system camera can be determined by:

- *Stadimetry,* also known as direct imaging, this is a technique for measuring distance based on the apparent size of an object in the field of view of the camera (Fig. 7a). The farther away the object, the smaller will be its apparent image. This approach requires an accurate focusing of the image.

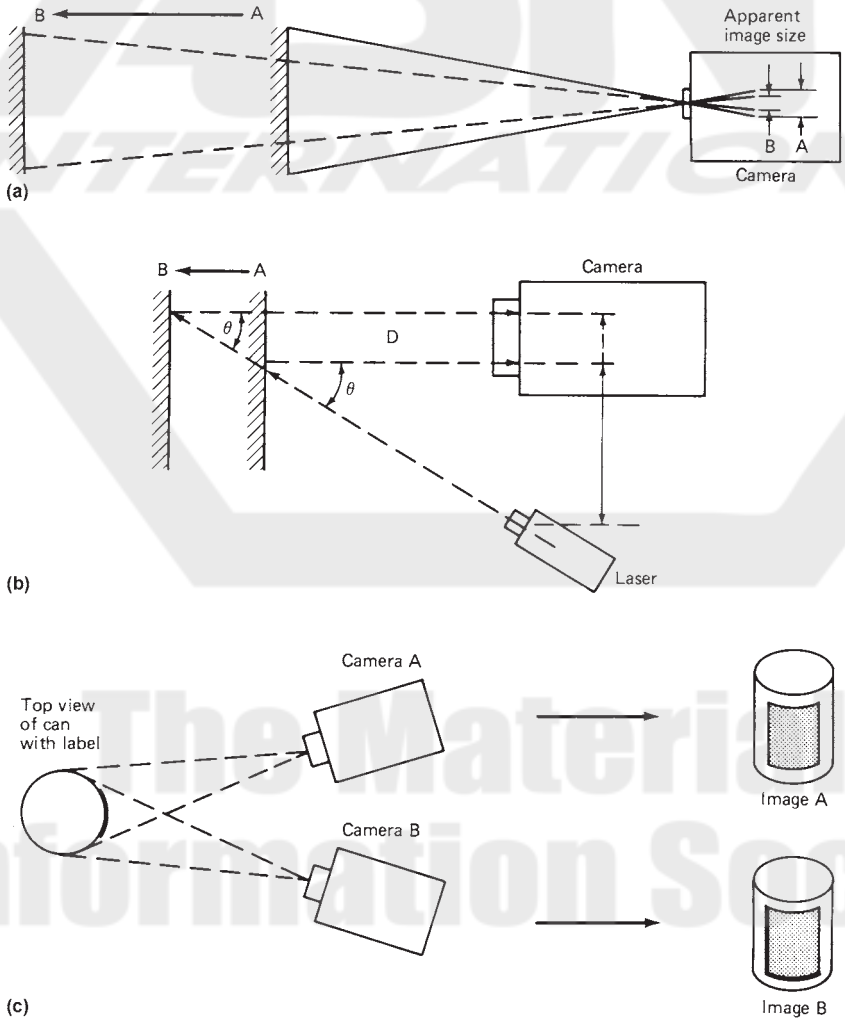


Fig. 7 Techniques for measuring the distance of an object from a vision system camera. (a) Stadimetry. (b) Triangulation. (c) Stereo vision. Source: Ref 1

- *Triangulation* is based on the measurement of the baseline of a right triangle formed by the light path to the object, the reflected light path to the camera, and a line from the camera to the light source (Fig. 7b). A typical light source for this technique is an LED or a laser, both of which form a well defined spot of light. Because the angle between the two light paths is preset, the standoff distance is readily calculated. Typical accuracies of 1 μm (40 μin.) can be achieved.
- *Stereo vision*, also known as binocular vision, is a method that uses the principle of parallax to measure distance (Fig. 7c). Parallax is the change in the relative perspective of a scene as the observer (or camera) moves. Human eyesight provides the best example of stereo vision. The right eye views an object as if the object were rotated slightly from the position observed by the left eye. Also, an object in front of another object seems to move relative to the other object when seen from one eye and then from the other. The closer the objects, the greater the parallax.

Object orientation is important in manufacturing operations such as material handling or assembly to determine where a robot may need to position itself relative to a part to grasp the part and then transfer it to another location. Among the methods used for determining object orientation are:

- *Equivalent Ellipse.* For an image of an object in a two-dimensional plane, an ellipse can be calculated that has the same area as the image. The major axis of the ellipse will define the orientation of the object. Another similar measure is the axis that yields the minimum moment of inertia of the object.
- *Connecting of Three Points.* If the relative positions of three noncolinear points in a surface are known, the orientation of the surface in space can be determined by measuring the apparent relative position of the points in the image.
- *Light Intensity Distribution.* A surface will appear darker if it is oriented at an angle other than normal to the light source. Determining orientation based on relative light intensity requires knowledge of the source of illumination as well as the surface characteristics of the object.
- *Structured light* involves the use of a light pattern rather than a diffused light source. The workpiece is illuminated by the structured light, and the way in which the pattern is distorted by the part can be used to determine both the three-dimensional shape and the orientation of the part.

Object Position Defined by Relative Motion. Certain operations, such as tracking or part insertion, may require the vision system to follow

the motion of an object. This is a difficult task that requires a series of image frames to be compared for relative changes in position during specified time intervals. Motion in one-dimension, as in the case of a moving conveyor of parts, is the least complicated motion to detect. In two-dimensions, motion may consist of both a rotational and a translational component. In three-dimensions, a total of six motion components (three rotational axes and three translational axes) may need to be defined.

Feature Extraction. One of the useful approaches to image interpretation is analysis of the fundamental geometric properties of two-dimensional images. Parts tend to have distinct shapes that can be recognized on the basis of elementary features. These distinguishing features are often simple enough to allow identification independent of the orientation of the part. For example, if surface area (number of pixels) is the only feature needed for differentiating the parts, then orientation of the part is not important. For more complex three-dimensional objects, additional geometric properties may need to be determined, including descriptions of various image segments. The process of defining these elementary properties of the image is often referred to as feature extraction. The first step is to determine boundary locations and to segment the image into distinct regions. Next, certain geometric properties of these regions are determined. Finally, these image regions are organized in a structure describing their relationship.

Light Intensity Variations. One of the most sophisticated and potentially useful approaches to machine vision is the interpretation of an image based on the different intensity of light in different regions. Many of the features described above are used in vision systems to create two-dimensional interpretations of images. However, analysis of subtle changes in shadings over the image can add a great deal of information about the three-dimensional nature of the object.

The problem is that most machine vision techniques are not capable of dealing with the complex patterns formed by varying conditions of illuminations, surface texture and color, and surface orientation. Another, more fundamental difficulty is that image intensities can change drastically with relatively modest variations in illumination or surface condition. Systems that attempt to match the gray level values of each pixel to a stored model can easily suffer deterioration in performance in real world manufacturing environments. The use of such geometric features such as edges or boundaries is likely to remain the preferred approach. Even better approaches are likely to result from research being performed on various techniques for determining surface shapes from relative intensity levels. For example, one approach assumes that the light intensity at a given point on the surface of an object can be precisely determined by an equation describing the nature and location of the light source, the orientation of the surface at the point, and the reflectivity of the surface.

Image Interpretation

When the system has completed the process of analyzing image features, the fourth step is performed. Some conclusions must be drawn with regard to the findings, such as the verification that a part is or is not present; the identification of an object based on recognition of its image; or the determination that certain parameters of the object fall within acceptable limits. Based on these conclusions, certain decisions can then be made regarding the object or the production process. These conclusions are formed by comparing the results of the analysis with a prestored set of standard criteria. These standard criteria describe the expected characteristics of the image and are developed either through a programmed model of the image or by building an average profile of previously examined objects.

Statistical Approach. In the simplest case of a binary system, the process of comparing an image with standard criteria, may simply require that all white and black pixels within a certain area be counted. Once the image is segmented or windowed, all groups of black pixels within each segment that are connected (called *blobs*) are identified and counted. The same process is followed for groups of white pixels (called *holes*).

The blobs, holes, and pixels are counted and the total quantity is compared with expected numbers to determine how closely the real image matches the standard image. If the numbers are within a certain percentage of each other, it can be assumed that there is a match.

An example of the statistical approach to image interpretation is the identification of a part on the basis of a known outline, such as the center hole of a washer. As illustrated in Fig. 8, a sample 3×3 pixel window can be used to locate the hole of the washer and to distinguish the washer from other distinctly different washers. The dark pixels shown in Fig. 8(a) represents the rough shape of the washer. When the window is centered on the shape, all nine white pixels are assigned a value of 0. In Fig. 8(b), a defective washer appears, with the hole skewed to the right. Because the window counts only six white pixels, it can be assumed that the hole is incorrectly formed. In Fig. 8(c), a second washer category is introduced, one with a smaller hole. In this case, only five white pixels are counted, providing enough information to identify the washer as a different type. In Fig. 8(d), a third washer is inspected, one that is larger than the first. The window counts nine white pixels, as in Fig. 8(a). In this case, some ambiguity remains, and so additional information would be required, such as the use of a 5×5 window. Another approach is to count all the black pixels rather than the white ones.

Such simple methods are finding useful applications in manufacturing because of the controlled, structured nature of most manufacturing environments. The extent of the analysis required for part recognition depends on both the complexity of the image and the goal of the analysis. In a

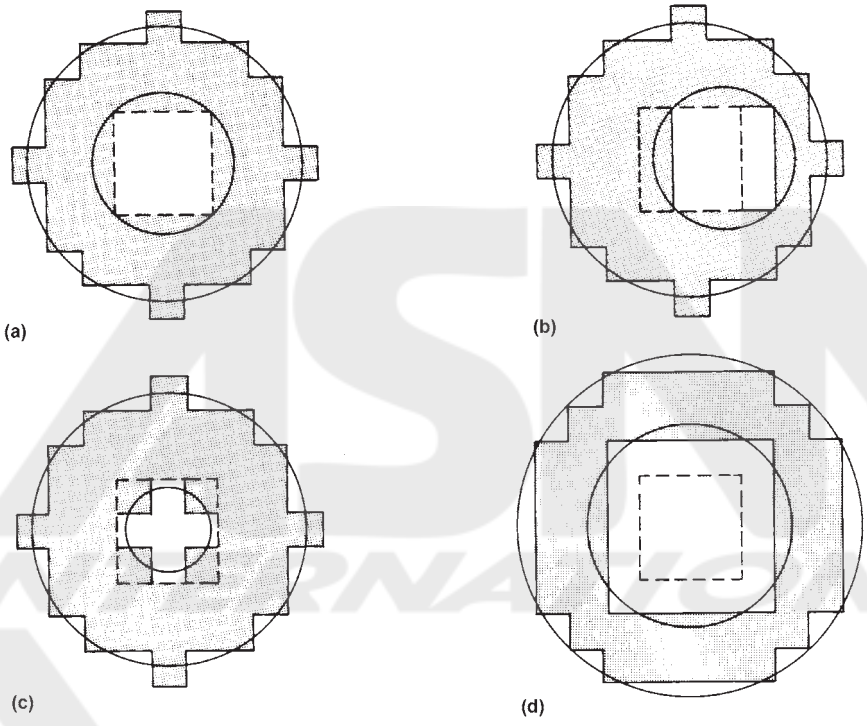


Fig. 8 Examples of the binary interpretation of washers using windowing. (a) Standard washer (9 pixels in window). (b) Washer with off-center hole (6 pixels in window). (c) Washer with small hole (5 white pixels in window). (d) Large washer (9 white pixels in window; need larger window). Source: Ref 1

manufacturing situation, the complexity of the image is often greatly reduced by controlling such factors as illumination, part location, and part orientation. The goal of the analysis is simplified when parts have easily identifiable features, as in the example of the washer.

Gray Scale Image Interpretation Versus Algorithms. There are two general ways in which image interpretation capabilities are being improved in vision systems. The first is gray scale image interpretation, and the second is the use of various algorithms for the complex analysis of image data. The use of gray scale image analysis greatly increases the quality of the data available for interpreting an image. The use of advanced data analysis algorithms (see Table 3) improves the way in which the data are interpreted. Both of these approaches allow the interpretation of much more complex images than the simple washer inspection example. However, even gray scale image analysis and sophisticated data analysis algorithms do not provide absolute interpretation of images.

Machine vision deals in probabilities, and the goal is to achieve a probability of correct interpretation as close to 100% as possible. In complex situations, human vision is vastly superior to machine systems. However,

Table 3 Typical software library of object location and recognition algorithms available from one machine vision system supplier

Tool	Function	Applications
Search	Locates complex objects and features	Fine alignment, inspection, gaging, guidance
Auto-train	Automatically selects alignment targets	Wafer and PCB alignment without operator involvement
Scene angle finder	Measures angle of dominant linear patterns	Coarse object alignment, measuring code angle for reading
Polar coordinate vision	Measures angle; handles circular images	Locating unoriented parts, inspecting and reading circular parts
Inspect	Performs Standford Research Institute (SRI) feature extraction (blob analysis)	Locating unoriented parts, defect analysis, sorting, inspection
Histograms	Calculates intensity profile	Presence/absence detection, simple inspection
Projection tools	Collapses 2-dimensional images into 1-dimensional images	Simple gaging and object finding
Character recognition	Reads and verifies alphanumeric codes	Part tracking, date/lot code verification
Image processing library	Filters and transforms images	Image enhancement, rotation, background filtering
V compiler	Compiles C language functions incrementally	All
Programming utilities	Handles errors, aids debugging	All
System utilities	Acquires images, outputs results, draws graphics	All
C library	Performs mathematics, creates reports and menus	All

Source: Ref 3

in many simple manufacturing operations, where inspection is performed over long periods of time, the overall percentage of correct conclusions can be higher for machines than for humans, who are subject to fatigue.

The two most commonly used methods of interpreting images are:

- *Feature Weighting.* In cases in which several image features must be measured to interpret an image, a simple factor weighting method can be used to consider the relative contribution of each feature to the analysis. For example, the image area alone may not be sufficient to ensure the positive identification of a particular valve stem in a group of valve stems of various sizes. The measurement of height and the determination of the centroid of the image may add some additional information. Each feature would be compared with a standard for a goodness-of-fit measurement. Features that are known to be the most likely indicators of a match would be weighted more than others. A weighted total goodness-of-fit score could then be determined to indicate the likelihood that the object has been correctly identified.
- *Template Matching.* In this method, a mask is electronically generated to match a standard image of an object. When the system inspects other objects in an attempt to recognize them, it aligns the image of each object with that of the standard object. In the case of a perfect match, all pixels would align perfectly. If the objects are not precisely the same, some pixels will fall outside of the standard image. The

percentage of pixels in two images that match is a measure of the goodness-of-fit. A threshold value can then be assigned to test for pass (positive match) or reject (no match) mode. A probability factor, which presents the degree of confidence with which a correct interpretation has been made, is normally calculated, along with the go/no-go conclusions.

Variations on these two approaches are used in most commercially available vision systems. Although conceptually simple, they can yield powerful results in a variety of manufacturing applications requiring the identification of two-dimensional parts with well defined silhouettes. With either method, a preliminary session is usually held before the machine is put into use. During this session, several sample known parts are presented to the machine for analysis. The part features are stored and updated as each part is presented, until the machine is familiar with the part. Then, the actual production parts are studied by comparison with this stored model of a standard part.

Mathematical Modeling. Although model building, or programming, is generally accomplished by presenting a known sample object to the machine for analysis, it is also possible to create a mathematical model describing the expected image. This is generally applicable for objects that have well defined shapes, such as rectangles or circles, especially if the descriptive data already exist in an off-line data base for computer-aided design and manufacture (CAD/CAM). For example, the geometry of a rectangular machined part with several circular holes of known diameters and locations can be readily programmed. Because more complex shapes may be difficult to describe mathematically, it may be easier to teach the machine by allowing it to analyze a sample part. Most commercial systems include standard image processing software for calculating basic image features and comparing with models. However, custom programming for model generation can be designed either by the purchaser or by the vision system supplier. Off-line programming is likely to become increasingly popular as CAD/CAM interface methods improve.

Although the image interpretation techniques described apply to many, if not most, of the machine vision systems that are commercially available, there are still other approaches being used by some suppliers, particularly for special purpose systems for such applications as printed circuit board (PCB) inspection, weld seam tracking, robot guidance and control, and inspection of microelectronic devices and tooling. These special purpose systems often incorporate unique image analysis and interpretation techniques that exploit the constraints inherent in the applications. For example, some PCB inspection systems employ image analysis algorithms based on design rules rather than feature weighting or template matching. In the design rule approach, the inspection process is based on known

characteristics of a good product. For PCBs, this would include minimum conductor width and spacing between conductors. Also, each conductor should end with a solder pad if the board is correct. If these rules are not complied with, then the product is rejected.

Interfacing. A machine vision system will rarely be used without some form of interaction with other factory equipment, such as CAD/CAM devices, robots, or host computers. This interaction is the final element of the machine vision process, in which conclusions about the image are translated into actions. In some cases, the final action may take the form of cumulative storage of information in a host computer, such as counting the number of parts in various categories for inventory control. In other situations, a final action may be a specific motion, such as the transfer of parts into different conveyors, depending on their characteristics. Vision systems are being increasingly used for control purposes through the combination of vision systems and robots. In this case, the vision system greatly expands the flexibility of the robot.

For most applications, interfacing a machine vision system with other equipment is a straightforward task. Most systems are equipped with a number of input and output ports, including a standard RS232C interface. However, connecting a vision system to a robot is much more complicated because of timing constraints, data formats, and the inability of most robot controllers to handle vision system inputs. To overcome this problem, several robot and vision system manufacturers have developed integrated system capabilities.

Machine Vision Applications

Machine vision systems can be considered for use in most manufacturing applications in which human vision is currently required. Human vision is required for applications in which noncontact feedback is used to provide information about a production process or a part. For example, a human welder or machinist uses visual feedback to ensure that the correct relationship is maintained between the tool and the workpiece. Human assemblers visually analyze the position of parts so that other parts can be correctly aligned for insertion or some other form of mating. Quality control inspectors visually check products or parts to ensure that there are no defects, such as missing parts, damage, or incorrect location of various features.

As discussed previously, the primary strength of human vision is the ability to analyze qualitative aspects of an object or a scene. However, humans are not particularly adept at measuring quantitative data. For example, although human vision uses a sophisticated approach for depth perception that allows it to correctly determine the relative distance of an object, it is not able to measure a specific distance to an object other than

as a very rough estimate. In addition, human vision can measure dimensions only approximately. Humans must rely on some standard frame of reference for judging an object. A standard retained in the memory does not provide a very good frame of reference from which to make quantitative measurements. It is not absolute, and it will vary from individual to individual. Because humans are also subject to fatigue, the interpretation of a standard may change over time.

Machine vision systems are ideally suited to a number of applications in which their ability to interpret images consistently over long periods of time makes them perform better than humans. Machine vision systems are also beginning to be used in many new and unique applications that simply did not exist previously. This includes, for example, on-line inspections that were not economically feasible before and the use of machine vision to increase manufacturing flexibility and reduce dependence on expensive hard tooling. The net result is both improved product quality and lower production costs.

In deciding whether or not machine vision will be effective in a particular application, the user must consider the capabilities of machine vision versus the requirements of the application. Although many applications are suitable for automated vision sensing, there are several complex applications in which the sophisticated recognition capability of human vision is better, such as the inspection of certain complex three-dimensional objects. In general, machine vision systems are suitable for use in three categories of manufacturing applications:

- Visual inspection of a variety of parts, subassemblies, and finished products to ensure that certain standards are met
- Identification of parts by sorting them into groups
- Guidance and control applications, such as controlling the motion of a robot manipulator

Examples of each of these three areas are listed in Table 4.

Table 4 Typical applications of machine vision systems

Area	Applications	Area	Applications	Area	Applications
Visual inspection	Measurement of length, width, and area Measurement of hole diameter and position Inspection of part profile and contour Crack detection On-line inspection of assemblies Verification of part features Inspection of surface finish	Part identification	Optical character recognition Identification of parts for spray painting Conveyor belt part sorting Bin picking Keyboard and display verification	Guidance and control	Vision-assisted robot assembly Vision-assisted robot material handling Weld seam tracking Part orientation and alignment systems Determining part position and orientation Monitoring high-speed packaging equipment

Source: Ref 1

ACKNOWLEDGMENT

This chapter was adapted from Machine Vision and Robotic Inspection Systems by J.D. Meyer, *Nondestructive Evaluation and Quality Control*, Volume 17, *ASM Handbook*, 1992, p 29–45.

REFERENCES

1. "Machine Vision Systems: A Summary and Forecast," 2nd ed., Tech Tran Consultants, Inc., 1985
2. P. Dunbar, Machine Vision, *Byte*, Jan 1986
3. J.D. Meyer, Machine Vision and Robotic Inspection Systems, *Nondestructive Evaluation and Quality Control*, Vol 17, *ASM Handbook*, ASM International, 1992, p 29–45

CHAPTER 5

Hardness Testing

THE TERM *HARDNESS*, as it is used in industry, may be defined as the ability of a material to resist permanent indentation or deformation when in contact with an indenter under load. Generally, a hardness test consists of pressing an indenter of known geometry and mechanical properties into the test material. The hardness of the material is quantified using one of a variety of scales that directly or indirectly indicate the contact pressure involved in deforming the test surface. Because the indenter is pressed into the material during testing, hardness is also viewed as the ability of a material to resist compressive loads. The indenter may be spherical as in the Brinell test, pyramidal as in the Vickers and Knoop tests, or conical as in the Rockwell test. In the Brinell, Vickers, and Knoop tests, hardness value is the load supported by unit area of the indentation, expressed in kilograms per square millimeter (kgf/mm^2). In the Rockwell test, the depth of indentation at a prescribed load is determined and converted to a hardness number (without measurement units), which is inversely related to the depth.

Hardness testers can either be portable instruments or laboratory devices. Static indentation and rebound testing are discussed in this chapter. These two methods account for virtually all routine hardness testing in the metalworking industry. Static indentation hardness testing is the more widely used of the two methods, although rebound testing is extensively employed, particularly for hardness measurements on large workpieces or for applications in which visible or sharp impressions in the test surface cannot be tolerated.

Brinell Hardness Testing

The Brinell hardness test is simple and consists of applying a constant load, usually 500 to 3000 kg (1100 to 6600 lb), on a hardened steel ball

type indenter, 10 mm (0.4 in) in diameter, to the flat surface of a workpiece as shown in Fig. 1. The lower 500 kg load is usually used for testing nonferrous metals, such as copper and aluminum alloys, whereas the higher 3000 kg load is most often used for testing harder metals, such as steels and cast irons. The load is held for a specified time (10 to 15 seconds for iron or steel and about 30 seconds for softer metals), after which the diameter of the recovered indentation is measured in millimeters. This time period is required to ensure that plastic flow of the work metal has stopped.

Hardness is evaluated by taking the mean diameter of the indentation (two readings at right angles to each other) and calculating the Brinell hardness number (HB) by dividing the applied load by the surface area of the indentation according to the following formula:

$$\mathrm{HB} = \frac{2L}{\pi D\left(D - \sqrt{D^2 - d^2}\right)}$$

where L is the load in kilograms, D is the diameter of the ball in millimeters, and d is the diameter of the indentation in millimeters. However, it is not necessary to make the calculation for each test. Such calculations are available in table form for all diameters of indentations.

Highly hardened steel or other very hard metals cannot be tested by a hardened steel ball by the Brinell method, because the ball will flatten during penetration and a permanent deformation will take place. This problem is recognized in specifications for the Brinell tests. Tungsten carbide balls are recommended for Brinell testing materials of hardness from 444

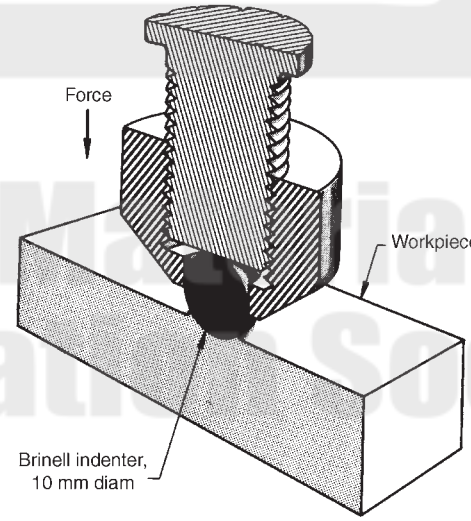


Fig. 1 Sectional view of a Brinell indenter, showing the manner in which the application of force by the indenter causes the metal of the workpiece to flow. Source: Ref 1

Brinell Hardness (HB) up to about 627 HB (indentation of 2.45 mm (0.095 in.) in diameter). However, higher Brinell values will be measured when using carbide balls instead of steel balls because of the difference in elastic properties. Therefore, the Brinell Hardness designation HBW is used when a tungsten carbide ball is used, and HBS is used when a hardened steel ball is used.

Surface Preparation. The degree of accuracy that can be attained by the Brinell hardness test can be greatly influenced by the surface smoothness of the test workpiece. The workpiece surface on which the Brinell indentation is to be made must be filed, ground, machined, or polished with emery paper (3/0 emery paper is suitable) so that the indentation diameter is clearly enough defined to permit its measurement. There should be no interference from tool marks.

Indentation Measurement. The diameter of the indentation is measured by a microscope to the nearest 0.05 mm (0.002 in.). This microscope contains a scale, and usually a built-in light, to facilitate easy reading.

The indentations produced in Brinell hardness tests may exhibit different surface characteristics. In some instances there is a ridge around the indentation that extends above the surface of the workpiece. In other instances the edge of the indentation is below the original surface. Sometimes there is no difference at all. The first phenomenon, called "ridging," is illustrated in Fig. 2(a). The second phenomenon, called "sinking," is illustrated in Fig. 2(b). An example of no difference is shown in Fig. 2(c). Cold-worked metals and decarburized steels are those most likely to exhibit ridging. Fully annealed metals and light case-hardened steels more often show sinking around the indentation.

Brinell Hardness Testers. Various kinds of Brinell testers are available for laboratory, production, automatic, and portable testing. These testers commonly use deadweight, hydraulic, pneumatic, elastic members (i.e.,

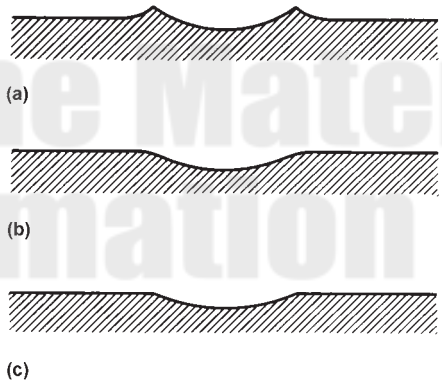


Fig. 2 Sectional views of Brinell indentations. (a) Ridging type Brinell impression. (b) Sinking type Brinell impression. (c) Flat type Brinell impression. Source: Ref 2

springs), or a closed loop load cell system to apply the test loads. All testers must have a rigid frame to maintain the load and a means of controlling the rate of load application to avoid errors due to impact (500 kgf/s or 1100 lbf/s maximum). The loads must be consistently applied within 1.0% as indicated in ASTM E10. In addition, the load must be applied so that the direction of load is perpendicular to the workpiece surface within two degrees for best results.

Bench units for laboratory testing are available with deadweight loading and/or pneumatic loading. Because of their high degree of accuracy, deadweight testers are most commonly used in laboratories and shops that do low to medium rate production. These units are constructed with weights connected mechanically to the Brinell ball indenter. Minimum maintenance is required because there are few moving parts. An example of a motorized deadweight tester is shown in Fig. 3(a).

Bench units are also available with pneumatic load application or a combination of deadweight/pneumatic loading. Figure 3(b) shows an example of the latter, where the load can be applied by release of deadweights or by pneumatic actuation. In both deadweight and pneumatic

Fig. 3 Bench type Brinell testers. (a) Motorized tester with deadweight loading. Courtesy of Wilson Instruments. (b) Brinell tester with combined deadweight loading and pneumatic operation. Courtesy of NewAge Industries. Source: Ref 2

bench units, the test piece is placed on the anvil, which is raised by an elevating screw until the test piece nearly touches the indenter ball. Operator controls initiate the load, which is applied at a controlled rate and time duration by the test machine. The test piece is then removed from the anvil, and the indentation width is measured with a Brinell scope, typically at 20× power. Testing with this type of apparatus is relatively slow and prone to operator influence on the test results.

Machines for Production Testing. Hydraulic testers were developed to reduce testing time and operator fatigue in production operations. Advantages of hydraulic testers include operating economy, simplicity of controls, and dependable accuracy. The controls prevent the operator from applying the load too quickly and thus overloading. The load is applied by a hydraulic cylinder and monitored by a pressure gage. Normally the pressure can be adjusted to apply any load between 500 and 3000 kgf (1100 and 6600 lbf). Hydraulic machines for production are available as bench top or as large floor units.

Automatic Testers. Many types of automatic Brinell testers are currently available. Most of these testers use a depth measurement system to eliminate the time consuming and operator biased measurement of the diameters. All of these testers use a preliminary load. Simple versions of this technique provide only comparative "go/no-go" hardness indications. More sophisticated models offer a microprocessor controlled digital readout to convert the depth measurement to Brinell numbers. Conversion from depth to diameter frequently varies for different materials and may require correlation studies to establish the proper relationship. These units can be fully automated to obtain production rates up to 600 tests per hour and can be incorporated into in line production equipment. The high speed automatic testers typically comply with ASTM E103, "Standard Method of Rapid Indentation Hardness Testing of Metallic Materials."

Portable Testing Machines. The use of conventional hardness testers may occasionally be limited because the work must be brought to the machine and because the workpieces must be placed between the anvil and the indenter. Portable Brinell testers that circumvent these limitations are available. A typical portable instrument is shown in Fig. 4. This type of tester weighs only about 11.4 kg (25 lb), so it can be easily transported to the workpiece. Portable testers can accommodate a wider variety of workpieces than can the stationary types. The tester attaches to the workpiece as would a C-clamp with the anvil on one side of the workpiece and the indenter on the other. For very large parts, an encircling chain is used to hold the tester in place as pressure is applied.

Portable testers generally apply the load hydraulically, employing a spring loaded relief valve. The load is applied by operating the hydraulic pump until the relief valve opens momentarily. With this type of tester, the hydraulic pressure should be applied three times when testing steel with a 3000 kgf load. This is equivalent to a holding time of 15 seconds, as re-

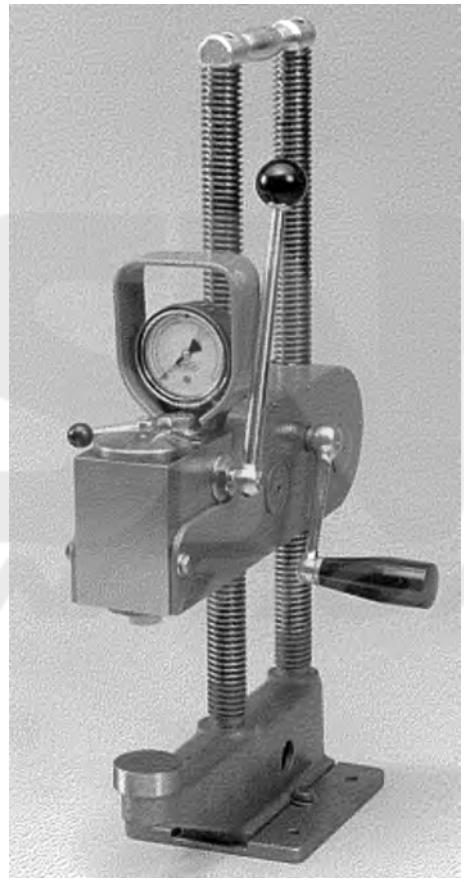


Fig. 4 Hydraulic, manually operated portable Brinell hardness tester. Source: Ref 2

quired by the more conventional method. For other materials and loads, comparison tests should be made to determine the number of load applications required to give results equivalent to the conventional method.

Spacing of Indentations. To ensure accurate results, indentations should not be made too close to the edge of the workpiece being tested. Lack of sufficient supporting material on one side of the workpiece will cause the resulting indentation to be large and unsymmetrical. It is generally agreed that the error in a Brinell hardness number is negligible if the distance from the center of the indentation is not less than 2.5 times (and preferably 3 times) the diameter of the indentation from any edge of the workpiece.

Similarly, indentations should not be made too close to one another. If indentations are too close together, the work metal may be cold worked by the first indentation, or there may not be sufficient supporting material for the second indentation. The latter condition would produce too large an indentation, whereas the former may produce too small an indentation. To

prevent this, the distance between centers of adjacent indentations should be at least three times the diameter of the indentation.

General Precautions. To avoid misapplication of Brinell hardness testing, the fundamentals and limitations of the test procedure must be clearly understood. Further, to avoid inaccuracies, some general rules to follow are:

- Indentations should not be made on a curved surface having a radius of less than 25 mm (1 in.)
- Spacing of indentations should be correct, as outlined in the section "Spacing of Indentations"
- The load should be applied steadily to avoid overloading caused by inertia of the weights
- The load should be applied in such a way that the direction of loading and the test surface are perpendicular to each other within two degrees
- The thickness of the workpiece being tested should be such that no bulge or mark showing the effect of the load appears on the side of the workpiece opposite the indentation. In any event, the thickness of the specimen shall be at least ten times the depth of indentation
- The surface finish of the workpiece being tested should be such that the indentation diameter is clearly outlined

Limitations. The Brinell hardness test has three principal limitations:

- Size and shape of the workpiece must be capable of accommodating the relatively large indentations
- Because of the relatively large indentations, the workpiece may not be usable after testing
- The limit of hardness range, about 11 HB with the 500 kg load to 627 HB with the 3000 kg (6600 lb) load, is generally considered the practical range

Rockwell Hardness Testing

Rockwell hardness testing is the most widely used method for determining hardness. There are several reasons for this distinction. The Rockwell test is simple to perform and does not require highly skilled operators. By use of different loads and indenters, Rockwell hardness testing can be used for determining hardness of most metals and alloys, ranging from the softest bearing materials to the hardest steels. A reading can be taken in a matter of seconds with conventional manual operation and in even less time with automated setups. No optical measurements are required; i.e., all readings are direct.

Rockwell hardness testing differs from Brinell hardness testing in that the hardness is determined by the depth of indentation made by a constant

load impressed upon an indenter. Although a number of different indenters are used for Rockwell hardness testing, the most common type is a diamond ground to a 120° cone with a spherical apex having a 0.2 mm (0.008 in.) radius, which is known as a Brale indenter, as depicted in Fig. 5(a).

The shape of the Rockwell diamond indenter most widely used in the United States is different from indenters used in the rest of the world. The ASTM specification calls for a diamond cone radius of 200 ± 10 μm (0.0079 in.), but in practice, it is closer to 192 μm (0.0076 in.). While not out-of-tolerance, the old U.S. standard indenter is at the low end of the specification. In the United States, the diamond was first set at 192 μm to match the nominal values of the hardness test blocks. However, the rest of the world has used a diamond size closer to 200 μm (0.0079 in.). A comparison of the old (192 μm) U.S. standard diamond indenter and the current (200 μm tip) U.S. indenter is shown in Fig. 5(b). The larger radius increases the indenter's resistance to penetration into the surface of the test piece.

At higher HRC hardness most of the indenter travel is along the radius; whereas at a lower HRC hardness, more indenter travel is along the angle. This is why the hardness shift from old to new has been most significant in the HRC 63 range and not the HRC 25 range.

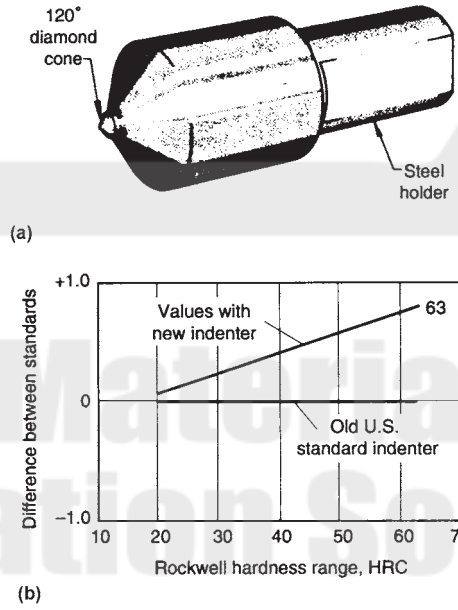


Fig. 5 Rockwell indenter. (a) Diamond cone Brale indenter (shown at about 2×). (b) Comparison of old and new U.S. diamond indenters. The angle of the new indenter remains at 120° but has a larger radius closer to the average ASTM specified value of 200 μm; the old indenter has a radius of 192 μm. The indenter with the larger radius has a greater resistance to penetration of the surface. Source: Ref 1

Rockwell Hardness Test Methods

As shown in Fig. 6, the Rockwell hardness test consists of measuring the additional depth to which an indenter is forced by a heavy (*major*) load beyond the depth of a previously applied light (*minor*) load. Application of the minor load eliminates backlash in the load train and causes the indenter to break through slight surface roughness and to crush particles of foreign matter, thus contributing to greater accuracy in the test. The basic principle involving minor and major loads illustrated in Fig. 6 applies to steel ball indenters as well as to diamond indenters.

The minor load is applied first, and a reference or *set position* is established on the measuring device of the Rockwell hardness tester. Then the major load is applied at a prescribed, controlled rate. Without moving the workpiece being tested, the major load is removed and the Rockwell hardness number is automatically indicated on the dial gage. The entire operation takes from 5 to 10 seconds.

Diamond indenters are used mainly for testing materials, such as hardened steels and cemented carbides. Steel ball indenters available with di-

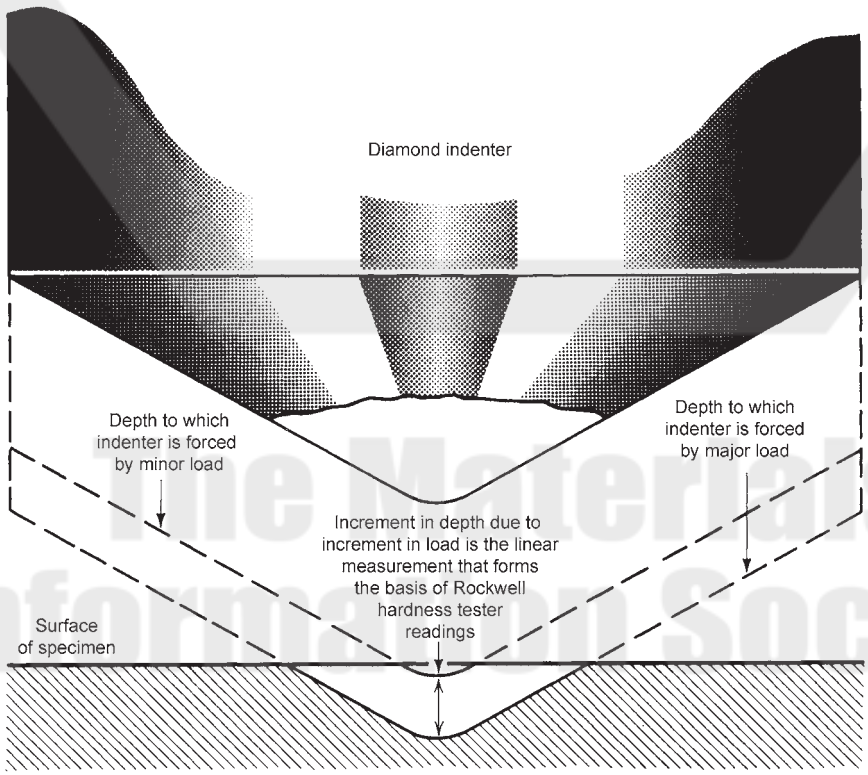


Fig. 6 Principle of the Rockwell test. Although a diamond indenter is illustrated, the same principle applies for steel ball indenters and other loads. Source: Ref 2

ameters of 1/16, 1/8, 1/4, and 1/2 inches, are used for testing materials, such as soft steel, copper alloys, aluminum alloys, and bearing metals.

Rockwell Testers. There are two basic types of Rockwell hardness testers—regular and superficial. Both testers have similar basic mechanical principles and significant components.

Rockwell testers generally come with two different resolutions. The standard Rockwell analog tester, shown in Fig. 7, that has been the industrial workhorse for years, has a resolution of 1.0 HRC. Many operators think they can improve resolution to 0.5 HRC or even 0.1 HRC by extrapolation, but this is not true. Extrapolation of readings only increases measurement error when several operators are checking parts. As with Brinell testing, better resolution can be achieved by investing in digital testing equipment. The newer digital Rockwell testers have a resolution of 0.1 HRC, and they eliminate the need for extrapolation or guessing. Similar resolution can be obtained on portable digital testers.

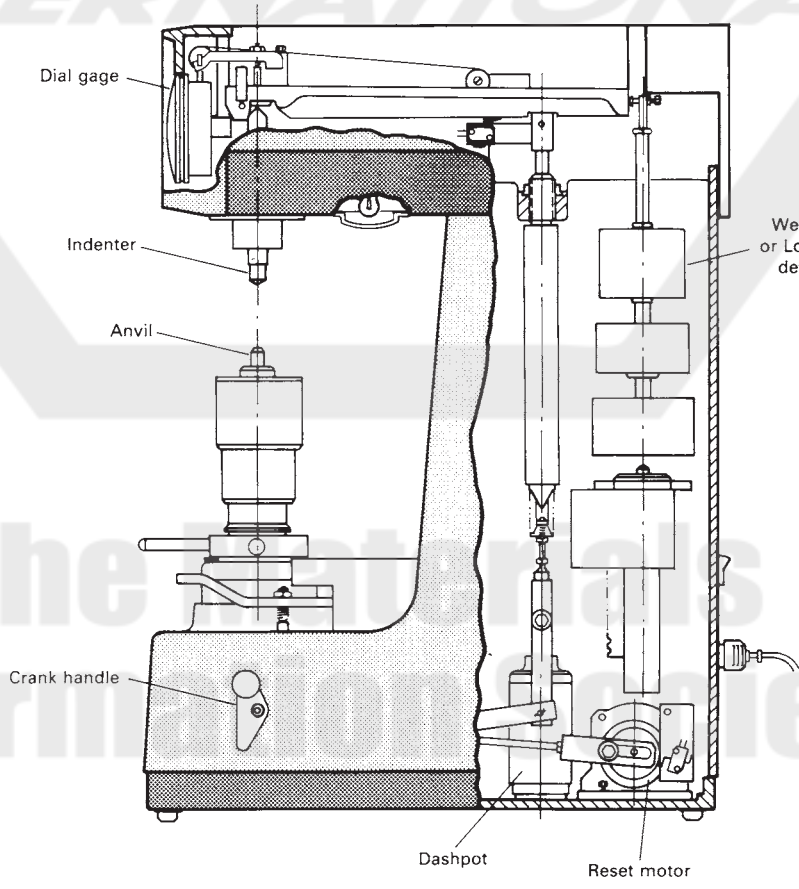


Fig. 7 Principal components of a regular or *normal* Rockwell hardness tester. Superficial Rockwell testers are similarly constructed. Source: Ref 2

Regular Rockwell Hardness Testing. In regular Rockwell hardness testing, the minor load is always 10 kg (22 lbs). The major load, however, can be 60, 100, or 150 kg (130, 220, or 330 lbs). No Rockwell hardness value is expressed by a number alone. A letter has been assigned to each combination of load and indenter, as shown in Table 1. Each number is suffixed by first the letter H (for hardness), then the letter R (for Rockwell), and finally the letter that indicates the scale used. For example, a value of 60 on the Rockwell C scale is expressed as 60 HRC, and so on. Regardless of the scale used, the set position is the same; however, when the diamond Brale indenter is used, the readings are taken from the black divisions on the dial gage. When testing with any of the ball indenters, the readings are taken from the red divisions.

One Rockwell number represents an indentation of 0.002 mm (0.00008 in.). Therefore, a reading of 60 HRC indicates indentation from minor to major load of (100 – 60) × 0.002 mm = 0.080 mm, or 0.0032 in. A reading of 80 HRB indicates an indentation of (130 – 80) × 0.002 mm = 0.100 mm, or 0.004 in.

Superficial Rockwell hardness testing employs a minor load of 3 kg (7 lb), but the major load can be 15, 30, or 45 kg (33, 66, or 99 lb). Just as in regular Rockwell testing, the indenter may either be a diamond or a steel ball, depending mainly on the nature of the metal being tested. Regardless of the load, the letter N designates use of the superficial Brale, and the letters T, W, X, and Y designate use of steel ball indenters. Scale and load combinations are presented in Table 1. Superficial Rockwell hardness values are always expressed with the number suffixed by a number and a letter that show the load/indenter combination. For example, if a load of 30 kg (66 lb) is used with a diamond indenter and a reading of 80

Table 1 Rockwell hardness scale designations for combinations of type of indenter and major load

Scale designation	Indenter		Major load, kg	Dial figure	Scale designation	Indenter		Major load, kg	Dial figure
	Type	Diam, in.				Type	Diam, in.		
Regular Rockwell tester					**Superficial Rockwell Tester**				
B	Ball	1/16	100	Red	15N	N Brale	...	15	...
C	Brale	...	150	Black	30N	N Brale	...	30	...
A	Brale	...	60	Black	45N	Brale	...	45	...
D	Brale	...	100	Black	15T	Ball	1/16	15	...
E	Ball	1/8	100	Red	30T	Ball	1/16	30	...
F	Ball	1/16	60	Red	45T	Ball	1/16	45	...
G	Ball	1/16	150	Red	15W	Ball	1/8	15	...
H	Ball	1/8	60	Red	30W	Ball	1/8	30	...
K	Ball	1/8	150	Red	45W	Ball	1/8	45	...
L	Ball	1/4	60	Red	15X	Ball	1/4	15	...
M	Ball	1/4	100	Red	30X	Ball	1/4	30	...
P	Ball	1/4	150	Red	45X	Ball	1/4	45	...
R	Ball	1/2	60	Red	15Y	Ball	1/2	15	...
S	Ball	1/2	100	Red	30Y	Ball	1/2	30	...
V	Ball	1/2	150	Red	45Y	Ball	1/2	45	...

Source: Ref 1

is obtained, the result is reported as 80 HR30N, where H means hardness, R means Rockwell, 30 means a load of 30 kg, and N indicates use of a diamond indenter.

All tests are started from the set position. One Rockwell superficial hardness number represents an indentation of 0.001 mm or 0.00004 in. Therefore, a reading of 80 HR30N indicates indentation from minor to major load of $(100 - 80) \times 0.001$ mm = 0.020 mm, or 0.0008 in. Dials on the superficial hardness testers contain only one set of divisions, which is used with all types of superficial indenters.

Portable Testing Machines. For hardness testing of large workpieces that cannot be moved, portable units are available in most regular and superficial scales and in a wide range of capacities (up to about a 355 mm, or 14 in., opening between anvil and indenter). Most portable hardness testers follow the Rockwell principle of minor and major loads, with the Rockwell hardness number indicated directly on the measuring device. Both digital and analog models are available. In Fig. 8(a), the workpiece is clamped in a C-clamp arrangement, and the indenter is recessed into a ring type holder that is part of the clamp. The test principal is identical to that of bench type models. The workpiece is held by the clamp between what is normally the anvil and the holder (which, in effect, serves as an upper anvil). The indenter is lowered to the workpiece through the holder. Other types of portable units (Fig. 8b) use the near-Rockwell method, where the diamond indenter is a truncated cone.

Selection of Rockwell Scale

Where no specification exists or there is doubt abut the suitability of a specified scale, an analysis should be made of those factors that influence

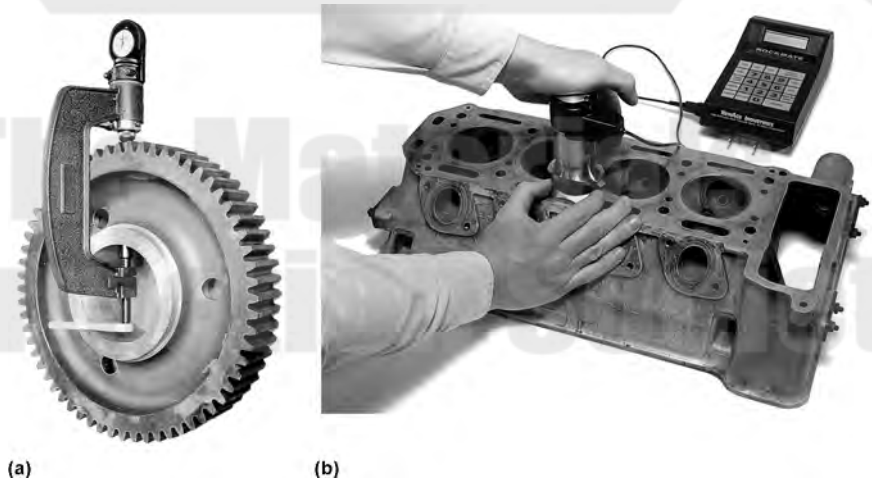


Fig. 8 Portable Rockwell testers. (a) C-clamp setup with a portable tester. (b) Portable near-Rockwell hardness tester. Source: Ref 2

the selection of the proper scale. These influencing factors are found in the following four broad categories:

- Type of work metal
- Thickness of work metal
- Width of area to be tested
- Scale limitation

Influence of Type of Work Metal. The types of work metal normally tested using the different regular Rockwell hardness scales are given in Table 2. This information also can be helpful when one of the superficial Rockwell scales may be required. For example, note that the C, A, and D scales, all with diamond indenters, are used on hard materials, such as steel and tungsten carbide. Any material in this hardness category would be tested with a diamond indenter. The choice to be made is whether the C, A, D, or the 45N, 30N, or 15N scale is applicable. Whatever the choice, the number of possible scales has been reduced to six. The next step is to find a scale, either regular or superficial, that will guarantee accuracy, sensitivity, and repeatability of testing.

Influence of Thickness of Work Metal. The metal immediately surrounding the indentation in a Rockwell hardness test is cold worked. The depth of material affected during testing is on the order of ten times the depth of the indentation. Therefore, unless the thickness of the metal being tested is at least ten times the depth of the indentation, an accurate Rockwell hardness test cannot be expected.

Influence of Test Area Width. In addition to the limitation of indentation depth for a workpiece of given thickness and hardness, there is a limiting factor on the minimum width of material. If the indentation is placed too close to the edge of a workpiece, the edge will deform outward and the Rockwell hardness number will be decreased accordingly. Experience has shown that the distance from the center of the indentation to the edge of

Table 2 Typical applications of regular Rockwell hardness scales

Scale(a)	Typical applications
B	Copper alloys, soft steels, aluminum alloys, malleable iron
C	Steel, hard cast irons, pearlitic malleable iron, titanium, deep case-hardened steel, and other materials harder than 100 HRB
A	Cemented carbides, thin steel, and shallow case-hardened steel
D	Thin steel and medium case-hardened steel and pearlitic malleable iron
E	Cast iron, aluminum and magnesium alloys, bearing metals
F	Annealed copper alloys, thin soft sheet metals
G	Phosphor bronze, beryllium copper, malleable irons. Upper limit is 92 HRG to avoid flattening of ball.
H	Aluminum, zinc, lead
K, L, M, P, R, S, V	Bearing metals and other very soft or thin materials. Use smallest ball and heaviest load that do not give anvil effect.

(a) The N scales of a superficial hardness tester are used for materials similar to those tested on the Rockwell C, A, and D scales but of thinner gage or case depth. The T scales are used for materials similar to those tested on the Rockwell B, F, and G scales but of thinner gage. When minute indentations are required, a superficial hardness tester should be used. The W, X, and Y scales are used for very soft materials . Source: Ref 1

the workpiece must be at least 2.5 times the diameter of the indentation to ensure an accurate test. Therefore, the width of a narrow test area must be at least 5 indentation diameters when the indentation is placed in the center.

Limitations of Rockwell Scales. The potential range of each Rockwell scale can be determined readily from the dial gage divisions on the tester: the black scale (for diamond indenter) on all regular hardness tester dial gages is numbered from 0 to 100, with 100 corresponding to the set position; the red scale (for ball indenters) is numbered from 0 to 130, with 130 being the set position. On the superficial hardness tester, the dial gage has only one set of divisions, numbered from 0 to 100.

Use of the diamond indenter when readings fall below 20 is not recommended, because there is loss of sensitivity when indenting this far down the conical section of the indenter. Brale indenters are not calibrated below values of 20, and if used on soft materials, there is no assurance that there will be the usual degree of agreement in results when replacing the indenters.

Support for Workpiece. A fundamental requirement of the Rockwell hardness test is that the surface of the workpiece being tested be approximately normal to the indenter and that the workpiece must not move or slip in the slightest degree as the major load is applied. The depth of indentation is measured by the movement of the plunger rod holding the indenter; therefore, any slipping or moving of the workpiece will be followed by the plunger rod and the motion transferred to the dial gage, causing an error to be introduced into the hardness test. As one point of hardness represents a depth of only 0.002 mm (0.00008 in.), a movement of only 0.025 mm (0.001 in.), could cause an error of over 10 Rockwell numbers. The support must be of sufficient rigidity to prevent its permanent deformation in use.

Anvils should be selected to minimize contact area of the workpiece while maintaining stability. Figure 9 illustrates several common types of anvils that can accommodate a broad range of workpiece shapes. An anvil with a large flat surface (Fig. 9b) should be used to support flat bottom workpieces of thick section. Anvils with a surface diameter greater than about 75 mm (3 in.) should be attached to the elevating screw by a threaded section, rather than inserted in the anvil hole in the elevating screw. Sheet metal and small workpieces that have flat undersurfaces are best tested on a spot anvil with a small, elevated, flat bearing surface (Fig. 9a). Workpieces that are not flat should have the convex side down on the bearing surface. Round workpieces should be supported in a V-slot anvil (Fig. 9a and c). Diamond spot anvils (Fig. 9d) are used only for testing very thin sheet metal samples in the HR15T and HR30T scales. Other anvil designs are available for a wide range of odd shaped parts, such as the eyeball anvil (Fig. 9e) that is used for tapered parts. Special anvils to accommodate specific workpiece configurations can be fabricated. Regardless of

anvil design, rigidity of the part to prevent movement during the test is absolutely essential for accurate results, as is cleanliness of the mating faces of the anvil and its supporting surface.

Work supports are available for long workpieces that cannot be firmly held on an anvil by the minor load. Because manual support is not practical, a jack rest should be provided at the overhang end to prevent pressure between the specimen and the penetrator. Figure 10 illustrates methods for testing long, heavy workpieces. When testing cylindrical pieces such as rods, the shallow-V or standard-V anvil should be used, and the indenter should be applied over the axis of the rod. Care should be taken that the specimen lies flat, supported by the sides of the V-slot anvil.

Fig. 9 Typical anvils for Rockwell hardness testing. (a) Standard spot, flat, and V anvils. (b) Testing table for large workpieces. (c) Cylinder anvil. (d) Diamond spot anvil. (e) Eyeball anvil. Source: Ref 2

Fig. 10 Rockwell test setups for long test pieces. (a) Jack setup. (b) Variable rest setup. Source: Ref 2

Vickers Hardness Testing (ASTM E384)

In 1925, Smith and Sandland of the United Kingdom developed a new indentation test for metals that were too hard to evaluate by the Brinell test, whose hardened steel ball was limited to steels with a hardness below ~450 HBS (~48 HRC). In designing the new indenter, they chose a square based diamond pyramid (Fig. 11) geometry that would produce hardness numbers nearly identical to Brinell numbers within the range of both. This decision was very wise, as it made the Vickers test very easy to adopt.

The ideal d/D ratio (d = impression diameter, D = ball diameter) for a spherical indenter is 0.375. If tangents are drawn to the ball at the impression edges for $d/D = 0.375$, they meet below the center of the impression at an angle of 136°, the angle chosen for the Vickers indenter.

The use of a diamond indenter allows the Vickers test to evaluate any material and, furthermore, has the very important advantage of placing the hardness of all materials on one continuous scale. The lack of a continuous scale is a major disadvantage of Rockwell type tests, for which 15 standard and 15 superficial scales were developed. Not one of these scales can cover the full hardness range. The HRA scale covers the broadest hardness range, but it is not commonly used.

In the Vickers test, the load is applied smoothly, without impact, and held in place for 10 or 15 seconds. The physical quality of the indenter and the accuracy of the applied load (defined in ASTM E384) must be controlled to get the correct results. After the load is removed, the two impression diagonals are measured, usually with a filar micrometer, to the nearest 0.1 μm, and then averaged. The Vickers hardness (HV) is calculated by:

$$\mathrm{HV} = \frac{1854.4L}{d^2}$$

where the load L is in grams-force, and the average diagonal d is in μm (although the hardness numbers are expressed in units of kgf/mm^2 rather than the equivalent gf/μm^2).

The original Vickers testers were developed for test loads of 1 to 120 kgf, which produce rather large indents. Recognizing the need for lower test loads, the National Physical Laboratory (U.K.) experimented with lower test loads in 1932. The first low load Vickers tester was described by Lips and Sack in 1936.

Because the shape of the Vickers indentation is geometrically similar at all test loads, the HV value is constant, within statistical precision, over a very wide test load range, as long as the test specimen is reasonably homogeneous. However, studies of microindentation hardness test results conducted over the past several years on a wide range of loads have shown that results are not constant at very low loads. This problem, called the *indentation size effect* (ISE), has been attributed to fundamental character-

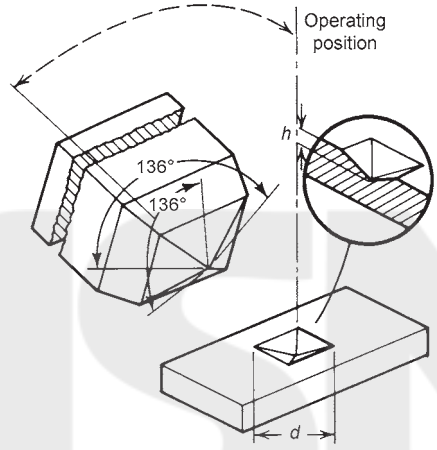


Fig. 11 Diamond pyramid indenter used for the Vickers test and resulting indentation in the workpiece. *d,* mean diagonal of the indentation in millimeters. Source: Ref 2

istics of the material. In fact, the same effect is observed at the low load test range of bulk Vickers testers.

Procedure. The Vickers hardness test follows the Brinell principle in that an indenter of definite shape is pressed into the material to be tested; the load removed; and, the diagonals of the resulting indentation measured. The indenter is made of diamond and is in the form of a square base pyramid having an angle of 136° between faces, as shown in Fig. 11. This indenter has angle across corners, or the so called *edge angle*, of 148° 6' 42.5''. The facets are highly polished and free from surface imperfections, and the point is sharp. The loads applied vary from 1 to 120 kg; the standard loads are 5, 10, 20, 30, 50, 100, and 120 kg. For most hardness testing, 50 kg is the maximum.

With the Vickers indenter, the depth of indentation is about one-seventh of the diagonal length of the indentation. For certain types of investigations, there are advantages to such a shape. The Vickers hardness number (HV) is the ratio of the load applied to the indenter to the surface area of the indentation. By formula:

$$\mathrm{HV} = \frac{2P\sin(\theta/2)}{d^2}$$

where P is the applied load in kilograms, d is the mean diagonal of the indentation in millimeters, and θ is the angle between opposite faces of the diamond indenter (136°).

Equipment for determining the Vickers hardness number should be designed to apply the load without impact, and friction should be reduced to a minimum. The actual load on the indenter should be correct to less than one percent, and the load should be applied slowly, because the Vickers is

a static test. Some standards require that the full load be maintained for 10 to 15 seconds. Loads of more than 50 kg (110 lb) are likely to fracture the diamond, especially when used on hard materials.

The accuracy of the micrometer microscope should be checked against a stage micrometer, which consists of ruled lines, usually 0.1 mm (0.004 in.) apart, that have been checked against certified length standards. The average length of the two diagonals is used in determining the hardness value.

The corners of the indentation provide indicators of the length of the diagonals. The area must be calculated from the average of readings of both diagonals. The indentations are usually measured under vertical illumination with a magnification of about 125 diameters.

The included angle of the diamond indenter should be 136° with a tolerance of less than ±0.50°, which is readily obtainable with modem diamond grinding equipment. This would mean an error of less than one percent in the hardness number. The indenters must be carefully controlled during manufacture so that in use the indentations produced will be symmetrical. Tables are available for converting the values of the diagonals of indentation in millimeters to the Vickers hardness number.

Vickers Hardness Testers. Several types of Vickers hardness testers are available. The principal component of a basic Vickers tester is shown schematically in Fig. 12(a), and a modern Vickers tester is shown in Fig. 12(b). Current equipment may include image analysis peripherals and other features for more automated handling and testing. Automation methods include motorized stage capabilities where long or repetitive hardness traverses are required. These can be programmed so that little operator involvement is required during the indentation mode. Digital display of the measured diagonals and automatic calculation of the hardness from the diagonals also simplify measurement but still requires the operator to peer into the microscope portion of the tester. Some systems include a closed circuit television system to the tester so that the operator can look at the magnified image on the TV screen and measure the diagonals. This is easier on the operators, but resolution of the system may not be as high.

Scleroscope Hardness Testing

The Scleroscope hardness test is essentially a dynamic indentation hardness test, wherein a diamond tipped hammer is dropped from a fixed height onto the surface of the material being tested. The height of rebound of the hammer is a measure of the hardness of the metal. The Scleroscope scale consists of units that are determined by dividing the average rebound of the hammer from a quenched (to maximum hardness) and untempered water hardening tool steel into 100 units. The scale is continued above 100 to permit testing of materials having a hardness greater than that of fully hardened tool steel.

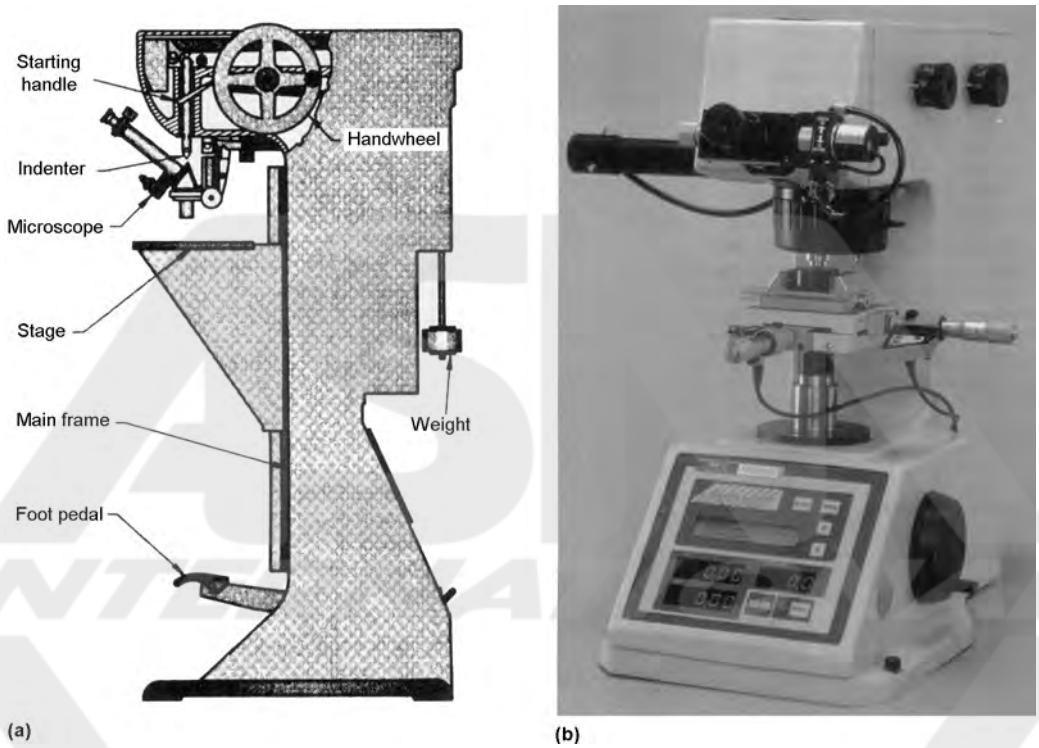


Fig. 12 Vickers hardness testers. (a) Principal components of a mechanical type. (b) Modern Vickers tester with digital readout of diagonal measurements and hardness values. Source: Ref 1

Testers. Two types of Scleroscope hardness testers are shown in Fig. 13. The Model C Scleroscope consists of a vertically disposed barrel containing a precision bore glass tube. A base mounted version of a Model C Scleroscope is shown in Fig. 13(a). The scale is graduated from 0 to 140. It is set behind and is visible through the glass tube. Hardness is read from the vertical scale, usually with the aid of the reading glass attached to the tester. A pneumatic actuating head, affixed to the top of the barrel, is manually operated by a rubber bulb and tube. The hammer drops and rebounds with the glass tube.

The Model D Scleroscope hardness tester (Fig. 13b) is a dial reading tester. The tester consists of a vertically disposed barrel that contains a clutch to arrest the hammer at maximum height of rebound, which is made possible because of the short rebound height. The hammer is longer and heavier than the hammer in the Model C Scleroscope and develops the same striking energy while dropping a shorter distance.

Both models of the Scleroscope hardness tester may be mounted on various types of bases. The C-frame base, which rests on three points and is for bench use in hardness testing small workpieces, has a capacity about 75 mm (3 in.) high by 64 mm (2.5 in.) deep. A swing arm and post is also for bench use but has height and reach capacities of 0.25 and 0.35 m (9

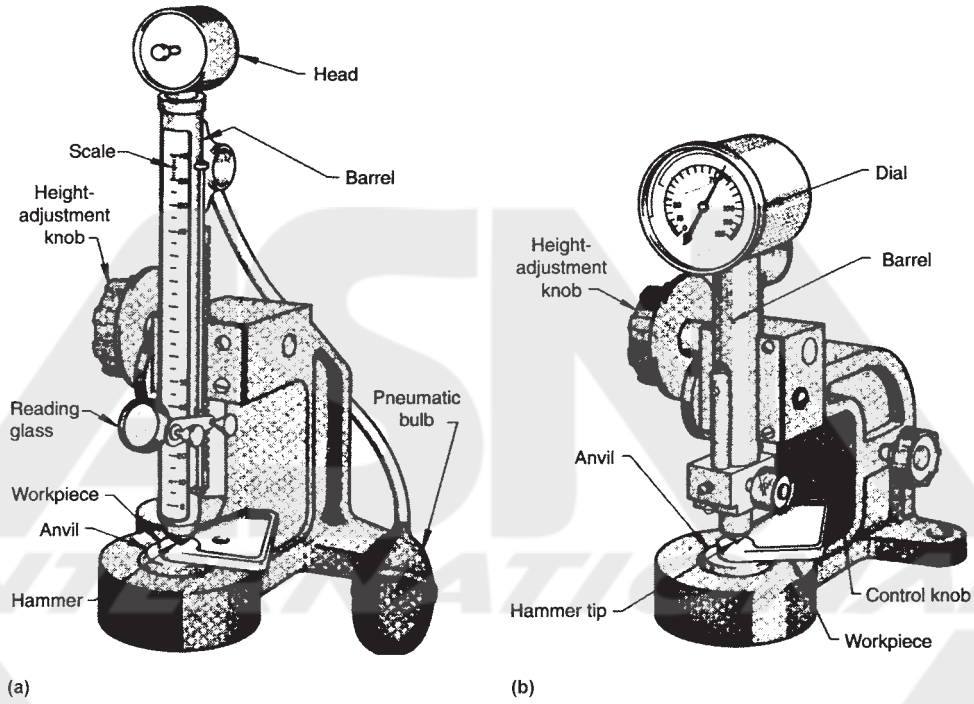


Fig. 13 Principal components of two types of base mounted Scleroscope hardness testers. (a) Model C, vertical scale. (b) Model D, dial reading. Source: Ref 1

and 14 in.), respectively. Another type of base is used for mounting the Scleroscope hardness tester on rolls and other cylindrical objects having a minimum diameter of 64 mm (2.5 in.), or on flat, horizontal surfaces having a minimum dimension of 75 by 130 mm (3 by 5 in.). The Model C Scleroscope hardness tester is commonly used unmounted. However, when the hardness tester is unmounted, the workpiece should have a minimum weight of 2.25 kg (5 lb). The Model D Scleroscope hardness tester should not be used unmounted.

Workpiece Surface Finish Requirements. As with other metallurgical hardness testers, certain surface finish requirements on the workpiece must be met for Scleroscope hardness testing to make an accurate hardness determination. An excessively coarse surface finish will yield erratic readings. Hence, when necessary, the surface of the workpiece should be filed, machined, ground, or polished to permit accurate, consistent readings to be obtained.

Limitations on Workpiece and Case Thickness. Case hardened steels having cases as thin as 0.25 mm (0.010 in.) can be accurately hardness tested provided the core hardness is no less than 30 Scleroscope. Softer cores require a minimum case thickness of 0.38 mm (0.015 in.) for accurate results.

Thin strip or sheet may be tested, with some limitations, but only when the Scleroscope hardness tester is mounted in the clamping stand. Ideally, the sheet should be flat and without undulation. If the sheet material is bowed, the concave side should be placed up to preclude any possibility of erroneous readings due to spring effect. The minimum thicknesses of sheet in various categories (in inches) that may be hardness tested are as follows:

Hardened steel	0.005
Cold finished steel strip	0.010
Annealed brass strip	0.015
Half-hard brass strip	0.010

Test Procedure. To perform a hardness test with either the Model C or the Model D Scleroscope hardness tester, the tester should be held or set in a vertical position, with the bottom of the barrel in firm contact with the workpiece. The hammer is raised to the elevated position and then allowed to fall and strike the surface of the workpiece. The height of rebound is then measured, which indicates the hardness. When using the Model C Scleroscope hardness tester, the hammer is raised to the elevated position by squeezing the pneumatic bulb. The hammer is released by again squeezing the bulb. When using the Model D Scleroscope hardness tester, the hammer is raised to the elevated position by turning the knurled control knob clockwise until a definite stop is reached. The hammer is allowed to strike the workpiece by releasing the control knob. The reading is recorded on the dial.

Spacing of Indentations. Indentations should be at least 0.50 mm (0.020 in.) apart and only one at the same spot. Flat workpieces with parallel surfaces may be hardness tested within 6 mm (0.25 in.) of the edge when properly clamped.

Taking the Readings. Experience is necessary to interpret the hardness readings accurately on a Model C Scleroscope hardness tester. Thin materials or those weighing less than 5 lb must be securely clamped to absorb the inertia of the hammer. The sound of the impact is an indication of the effectiveness of the clamping: a dull thud indicates that the workpiece has been clamped solid, whereas a hollow ringing sound indicates that the workpiece is not tightly clamped or is warped and not properly supported. Five hardness determinations should be made and their average taken as representative of the hardness of a particular workpiece.

Advantages. The advantages for using the Scleroscope hardness test are:

- Tests can be made very rapidly. Over 1000 tests per hour are possible
- Operation is simple and does not require highly skilled technicians

- The Model C Scleroscope tester is portable and may be used unmounted for hardness testing workpieces of unlimited size: rolls, large dies, and machine tool ways
- The Scleroscope hardness test is a nonmarring test; no crater is left, and only in the most unusual instances would the tiny hammer mark be objectionable on a finished workpiece
- A single scale accommodates the entire hardness range from the softest to the hardest metals

Limitations. The limitations when using the Scleroscope hardness test are:

- The hardness tester must be in a vertical position, or the free fall of the hammer will be impeded and result in erratic readings
- Scleroscope hardness tests are more sensitive to variations in surface conditions than some other hardness tests
- Because readings taken with the Model C Scleroscope hardness tester are those observed from the maximum rebound of the hammer on the first bounce, even the most experienced operators may disagree among themselves by one or two points in the reading

Microhardness Testing

The term *microhardness* usually refers to indentation hardness tests made with loads not exceeding 1 kg (2 lbs). Such hardness tests have been made with a load as light as 1 g (0.002 lbs), although the majority of microhardness tests are made with loads of 100 to 500 g (0.2 to 1 lb). In general, the term is related to the size of the indentation rather than the load applied. Microhardness testing is capable of providing information regarding the hardness characteristics of materials that cannot be obtained with hardness tests, such as the Brinell, Rockwell, or Scleroscope.

Because of the required degree of precision for both equipment and operation, microhardness testing is usually, although not necessarily, performed in a laboratory. However, such a laboratory is often a process control laboratory and may be located close to production operations. Microhardness testing is recognized as a valuable method for controlling numerous production operations in addition to its use in research applications. Specific fields of application of microhardness testing include:

- Measuring hardness of precision workpieces that are too small to be measured by the more common hardness testing methods
- Measuring hardness of product forms, such as foil or wire, that are too thin or too small in diameter to be measured by the more convenient methods

- Monitoring of carburizing or nitriding operations, which is usually accomplished by hardness surveys taken on cross sections of test pieces that accompanied the workpieces through production operations
- Measuring hardness of individual microconstituents
- Measuring hardness close to edges, thus detecting undesirable surface conditions, such as grinding bum and decarburization
- Measuring hardness of surface layers, such as plated or bonded layers

Indenters. Microhardness testing is performed with either the Knoop or the Vickers indenter. The Knoop indenter is more widely used in the United States, while the Vickers indenter is more widely used in Europe.

As previously discussed, the Vickers test is an indentation test that employs a square based pyramidal shaped indenter made from diamond (Fig. 14a). Figure 14(b) shows examples of Vickers indents to illustrate the influence of test force on indent size.

In this test, the force is applied smoothly, without impact, and held in contact for 10 to 15 seconds. The force must be known precisely (refer to ASTM E384 for tolerances). After the force is removed, both diagonals are measured and the average is used to calculate the HV according to:

$$\mathrm{HV} = \frac{2000P\sin\left(\alpha/2\right)}{d^2} = \frac{1854.4P}{d^2}$$

where d is the mean diagonal in μm, P is the applied load in gf, and α is the face angle (136°).

The hardness can be computed with the formula and a calculator, or using a spreadsheet program. Most modern microhardness test units have the calculation capability built in and display the hardness value along with the measured diagonals. A book of tables of HV as a function of d and P also accompanies most testers, and ASTM E384 includes such tables.

In 1939, Frederick Knoop and his associates at the former National Bureau of Standards developed an alternate indenter based on a rhombohedral shaped diamond with the long diagonal approximately seven times as long as the short diagonal (Fig. 15a). Figure 15(b) shows examples of Knoop indents to illustrate the influence of applied load on indent size. The Knoop indenter is used in the same machine as the Vickers indenter, and the test is conducted in exactly the same manner, except that the Knoop hardness (HK) is calculated based on the measurement of the long diagonal only and calculation of the projected area of the indent rather than the surface area of the indent:

$$\mathrm{HK} = \frac{1000P}{C_\mathrm{p}d^2} = \frac{14229P}{d^2}$$

where C_p is the indenter constant, which permits calculation of the projected area of the indent from the long diagonal squared.

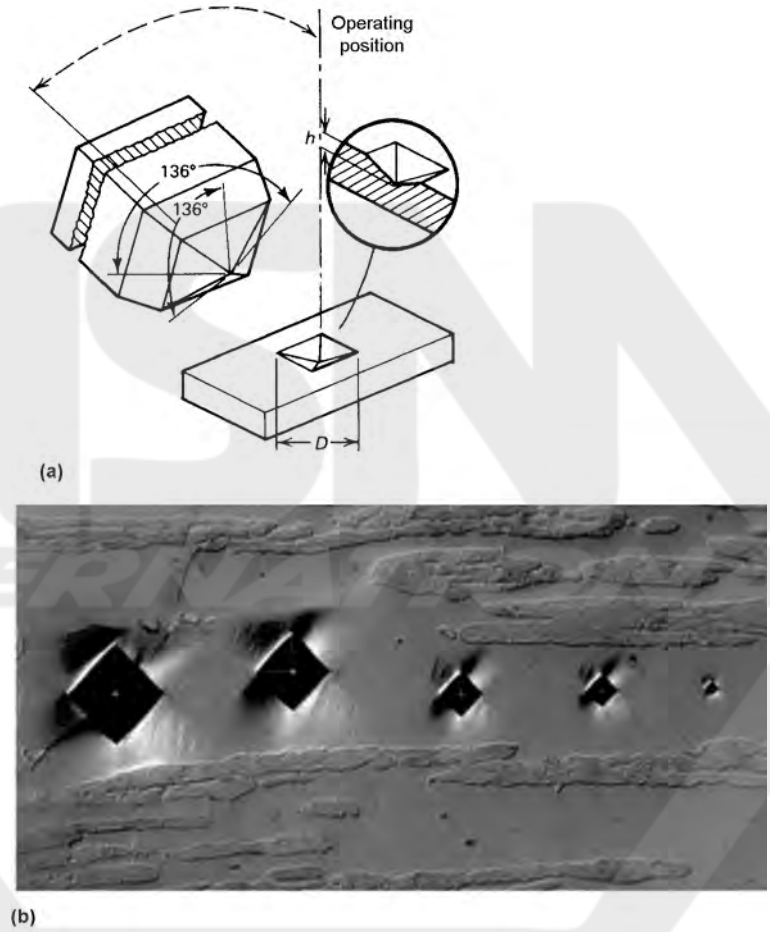


Fig. 14 Vickers hardness test. (a) Schematic of the square based diamond pyramidal indenter used for the Vickers test and an example of the indentation it produces. (b) Vickers indents made in ferrite in a ferritic martensitic high carbon version of 430 stainless steel using (left to right) 500, 300, 100, 50, and 10 gf test forces, differential interference contrast illumination, aqueous 60% nitric acid, 1.5 V dc. Original magnification 250×. Source: Ref 3

A comparison of the indentations made by the Knoop and Vickers indenters is shown in Fig. 16. Each has some advantages over the other. For example, the Vickers indenter penetrates about twice as far into the workpiece as does the Knoop indenter; and, the diagonal of the Vickers indentation is about one third of the total length of the Knoop indentation. Therefore, the Vickers indenter is less sensitive to minute differences in surface conditions than is the Knoop indenter. However, the Vickers indentation, because of the shorter diagonal, is more sensitive to errors in measurement than the Knoop indentation.

The shortcoming of the Knoop indent is that the three-dimensional indent shape changes with test load and, consequently, HK varies with load. In fact, HK values may be reliably converted to other test scales only for

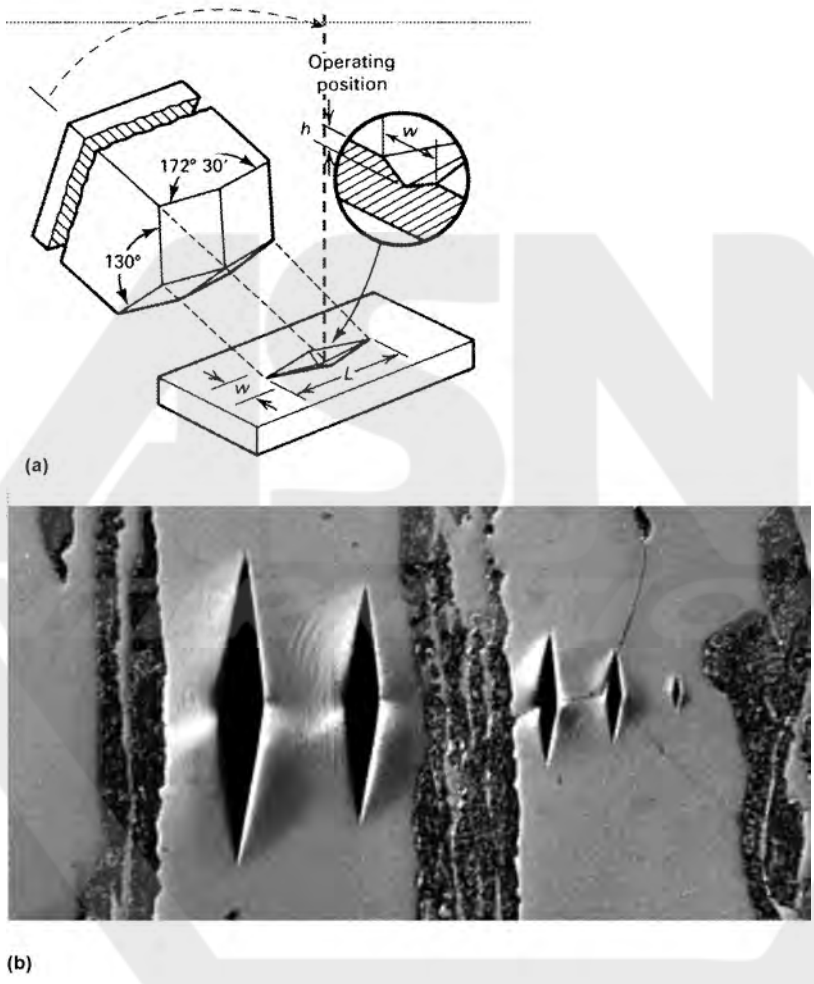


Fig. 15 Knoop hardness test. (a) Schematic of the rhombohedral shaped diamond indenter used for the Knoop test and an example of the indentation it produces. (b) Knoop indents made in ferrite in a ferritic martensitic high carbon version of 430 stainless steel using (left to right) 500, 300, 100, 50, and 10 gf test forces, differential interference contrast illumination, aqueous 60% nitric acid, 1.5 V dc. Original magnification 300×. Source: Ref 3

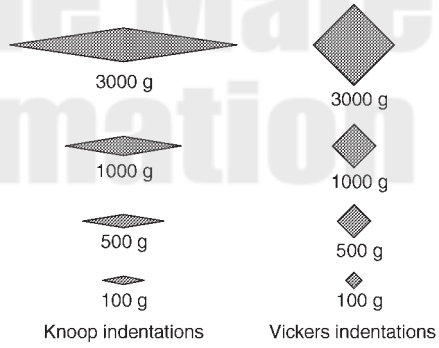


Fig. 16 Comparison of indentations made by Knoop and Vickers indenters in the same work metal and at the same loads. Source: Ref 1

HK values produced at the standard load, generally 500 gf, was used to develop the correlations. However, at high loads the variation is not substantial. Note that all hardness scale conversions are based on empirical data; consequently, conversions are not precise but are estimates.

Microhardness Testers. Several types of microhardness testers are available. The most accurate testers operate through the direct application of load by dead weight or by weights and lever.

The Tukon tester is widely used for microhardness testing. Several different designs of this microhardness tester are available; they vary mainly in load range, but all can accommodate both Knoop and Vickers indenters. The Tukon microhardness tester shown in Fig. 17 has a load range of 1 to 1000 g (0.002 to 2 lbs). Loads are applied by dead weight. The microscope is furnished with three objective lenses having magnifications of about 150, 300, and 600 diameters.

Sources of test error include inaccuracy in loading, vibration, rate of load application, duration of contact period, and impact. To limit the shock that can occur when the operator removes the load, this generally has an adverse effect on indentations made with loads below 500 g (1 lb), an au-

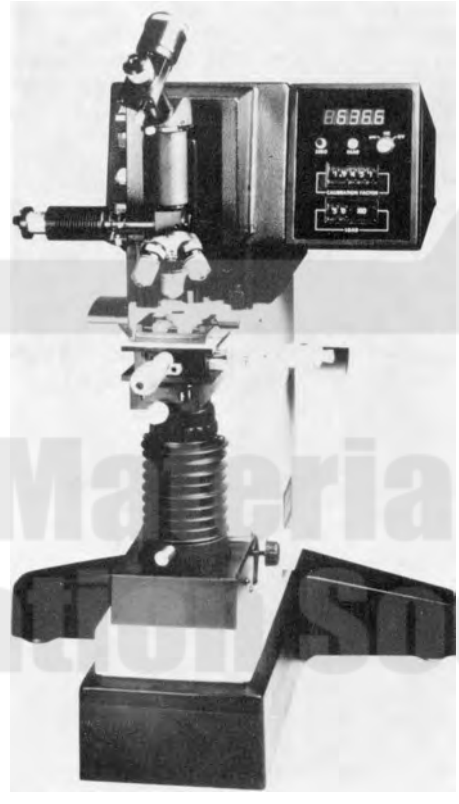


Fig. 17 Principal components of a Tukon microhardness tester. Source: Ref 1

tomatic test cycle is built into the Tukon microhardness tester. With this automatic test cycle, the load is applied at a constant rate, maintained in the work for 18 seconds, and smoothly removed. Thus, the operator does not need to touch the tester while the load is being applied and removed. The design of microhardness testers will vary from one type to another, but it is essential to remove the applied load without touching the tester if clear cut indentations are to be obtained.

A movable stage to support the workpiece is an essential component of a microhardness tester. In many applications the indentation must be in a selected area, usually limited to a few thousandths of a square millimeter. In testing with the type of Tukon microhardness tester shown in Fig. 17, first the required area is located by looking through the microscope and moving the mechanical stage until the desired location is centered within the optical field of view. The stage is then indexed under the indenter, and the automatic indentation cycle is initiated by tripping the handle. After the cycle is completed, signaled by a telltale light, the stage is again indexed back under the objective for indentation measurement.

Optical equipment used in microhardness testers for measuring the indentation must focus on both ends of the indentation at the same time, as well as be rigid and free from vibration. Lighting is also important. Complete specifications of measurement, including the mode of illumination, are necessary in microhardness testing techniques. Polarized light, for instance, results in larger measurements than does unpolarized light. Apparently, this is caused by the reversal of the diffraction pattern; that is, the indentation appears brighter than the background. When test data are recorded, it is recommended that both the magnification and the type of illumination used be reported.

In measuring the indentation, the proper illumination to obtain optimum resolution is essential, and the appropriate objective lens should be selected. In operation, the ends of the indentation diagonals should be brought into sharp focus. With the Knoop indenter, one leg of the long diagonal should not be more than 20% longer than the other. If this is not apparent or if the ends of the diagonal are not in focus, the surface of the workpiece should be checked to make sure it is normal to the axis of the indenter. With the Vickers indenter, both diagonals should be measured and the average used for calculating the Vickers hardness number (HV).

Preparing and Holding the Specimen. Regardless of whether the metal being tested for microhardness is an actual workpiece or a representative specimen, surface finish is of prime importance. To permit accurate measurement of the length of the Knoop indentation or diagonals of the Vickers indentation, the indentation must be clearly defined. In general, as the test load decreases, the surface finish requirements become more stringent. When the load is 100 g (0.2 lbs) or less, a metallographic finish is recommended.

Specific Applications of Microhardness Testing

Microhardness testing is used extensively in research and for controlling the quality of manufactured products, as well as for solving shop problems.

Testing of small workpieces is an important use of microhardness testing. Many manufactured products, notably in the instrument and electronics industries, are too small to be tested for hardness by the more conventional methods. Many such workpieces can be tested without impairing their usefulness, generally by means of various types of holding and clamping fixtures. Microhardness testing is also applied to product forms that cannot be tested by other means. Thin foils and small diameter wires are typical examples.

Monitoring of Surface Hardening Operations. Microhardness testing is the best method in present use for accurately determining case depth and certain case conditions of carburized or nitrided workpieces, using the hardness survey procedure. In most instances, this is accomplished by use of test coupons that have accompanied the actual workpiece through the heat treating operation. The coupons are then sectioned and usually mounted for testing. To ensure accurate readings close to the edge of the cross section, the 100 g (0.2 lbs) is most often used, although a 500 g (1 lb) load is sometimes preferred. If the 100 g load is used, a metallographic finish is essential. Readings are taken at pre-established intervals (commonly, 0.004 or 0.005 in.), usually beginning at least 0.001 in. from the edge of the workpiece.

Accuracy, Precision, and Bias. Many factors can influence the quality of microindentation test results, including instrument factors, measurement factors, and material factors:

- Instrument factors
  Accuracy of the applied load
  Inertia effects, speed of loading
  Angle of indentation
  Lateral movement of the indenter or specimen
  Indentation time
  Indenter shape deviations
  Damage to the indenter
  Insufficient spacing between indents or from edges
- Measurement factors
  Calibration of the measurement system
  Resolving power of the objective
  Magnification
  Operator bias in sizing
  Inadequate image quality
  Nonuniform illumination

- Material factors
  Heterogeneity in composition or microstructure
  Crystallographic texture
  Quality of the specimen preparation
  Low reflectivity or transparency

In the early days of low load (<100 gf or < 0.2 lbf) hardness testing, it was quickly recognized that improper specimen preparation can influence hardness test results. Most texts state that improper preparation yields higher test results because the surface contains excessive preparation induced deformation. While this is certainly true, improper preparation may also create excessive heat, which reduces the hardness and strength of many metals and alloys. Either problem may be encountered due to faulty preparation.

For many years, it was considered necessary to electrolytically polish specimens so that the preparation induced damage could be removed, thus permitting bias free low load testing. However, the science behind mechanical specimen preparation, chiefly due to the work of L. Samuels, has led to development of excellent mechanical specimen preparation procedures, and electropolishing is no longer required.

In addition, several operational factors must be controlled for optimum test results. First, it is good practice to inspect the indenter periodically for damage, for example, cracking or chipping of the diamond. If you have metrology equipment, you can measure the face angles and the sharpness of the tip. Specifications for Vickers and Knoop indenter geometries are given in ASTM E384.

A prime source of error is the alignment of the specimen surface relative to the indenter. The indenter itself must be properly aligned perpendicular (±1°) to the stage plate. Next, the specimen surface must be perpendicular to the indenter. Most testers provide holders that align the polished face perpendicular to the indenter or parallel to the stage. If a specimen is simply placed on the stage surface, its back surface must be parallel to its polished surface. Tilting the surface more than one degree perpendicular to the indenter results in nonsymmetrical impressions and can produce lateral movement between specimen and indenter. However, in most cases, indenting procedures are not the major source of error.

It is important to regularly check the performance of the tester with a certified test block. The safest choice is a test block manufactured for microindentation testing and certified for the test (Vickers or Knoop) as well as the specified load. Strictly speaking, a block certified for Vickers testing at 300 or 500 gf (0.7 or 1 lbf), commonly chosen loads, should yield essentially the same hardness with loads from about 50 to 1000 gf (0.1 to 2 lbf). That is, if you take the average of about five indents and compare the average at test load to the average at the calibrated load, knowing the

standard deviation of the test results, statistical tests can tell at any desired confidence level, if the difference between the mean values of the tests at the two loads is statistically significant or not.

When considering a new tester, it is prudent to perform a series of indents, five is adequate, at each test load (L) available. Then, plot the mean and 95% confidence limits of each test as a function of load. Because of the method of defining HV and HK, which involves dividing by d^2, measurement errors become more critical as d gets smaller; that is, as L decreases and the material's hardness increases. Therefore, departure from a constant hardness for the Vickers or Knoop tests as a function of load becomes a greater problem as the hardness increases. For the Knoop test, HK increases as L decreases because the indent geometry changes with indent depth and width. But the change in HK varies with the test load. At a higher test load the change is greater as L decreases.

The greatest source of error is measuring the indent. The indent should be placed in the center of the measuring field, because lens image quality is best in the center. The light source should provide adequate, even illumination to provide maximum contrast and resolution. The accuracy of the filar micrometer or other measuring device should be verified by a stage micrometer.

Specimen preparation quality becomes more important as the load decreases, and it must be at an acceptable level. Specimen thickness must be at least 2.5 times the Vickers diagonal length. Because the Knoop indent is shallower than the Vickers at the same load, somewhat thinner specimens can be tested. Spacing of indents is important because indenting produces plastic deformation and a strain field around the indent. If the spacing is too small, the new indent will be affected by the strain field around the last indent. ASTM recommends a minimum spacing (center to edge of adjacent indent) of 2.5 times the Vickers diagonal. For the Knoop test, in which the long diagonals are parallel, the spacing is 2.5 times the short diagonal. The minimum recommended spacing between the edge of the specimen and the center of the indent should be 2.5 times. Again, Knoop indents can be placed closer to the surface than Vickers indents.

ACKNOWLEDGMENT

This chapter was adapted from Hardness Testing, *Metals Handbook Desk Edition*, Second Edition, 1998, and Macroindentation Hardness Testing by E.L. Tobolski and A. Fee in *Mechanical Testing and Evaluation*, Volume 8, *ASM Handbook*, 2000.

REFERENCES

1. Hardness Testing, *Metals Handbook Desk Edition*, 2nd ed., ASM International, 1998, p 1308–1317

2. E.L. Tobolski and A. Fee, Macroindentation Hardness Testing, *Mechanical Testing and Evaluation*, Vol 8, *ASM Handbook*, ASM International, 2000, p 203–220
3. G.F. Vander Voort, Microindentation Hardness Testing, *Mechanical Testing and Evaluation*, Vol 8, *ASM Handbook*, ASM International, 2000, p 221–231

SELECTED REFERENCES

- *Mechanical Testing and Evaluation*, Vol 8, *ASM Handbook*, ASM International, 2000

Brinell Hardness Standards for Metals

- ASTM E10, Standard Test Method for Brinell Hardness of Metallic Materials
- BS EN ISO 6506-1, Metallic Materials—Brinell Hardness Test—Test Method
- BS EN ISO 6506-2, Metallic Materials—Brinell Hardness Test—Verification and Calibration of Brinell Hardness Testing Machines
- BS EN ISO 6506-3, Metallic Materials—Brinell Hardness Test—Calibration of Reference Blocks
- DIN EN, Brinell Hardness Test—Test Method 10003-1
- DIN EN 10003-2, Metallic Materials—Brinell Hardness Test—Verification of Brinell Hardness Testing Machines
- DIN EN 10003-3, Metallic Materials—Brinell Hardness Test—Calibration of Standardized Blocks to be Used for Brinell Hardness Testing Machines
- JIS B 7724, Brinell Hardness Testing Machines
- JIS B 7736, Standardized Blocks of Brinell Hardness
- JIS Z 2243, Method of Brinell Hardness Test

Rockwell Hardness Standards for Metals

- ASTM B294, Standard Test Method for Hardness Testing of Cemented Carbides
- ASTM E18, Test Methods for Hardness and Rockwell Superficial Hardness of Metallic Materials
- ASTM E1842, Test Method for Macro-Rockwell Hardness Testing of Metallic Materials
- BS 5600-4.5, Powder Metallurgical Materials and Products—Methods of Testing and Chemical Analysis of Hardmetals—Rockwell Hardness Test (Scale A)
- BS EN ISO 6508-1, Metallic Materials—Rockwell Hardness Test—Part 1: Test Method (Scales A, B, C, D, E, F, G, H, K, N, T)
- BS EN ISO 6508-2, Metallic Materials—Rockwell Hardness Test—Part 2: Verification and Calibration of Testing Machines (Scales A, B, C, D, E, F, G, H, K, N, T)

- BS EN ISO 6508-3, Metallic Materials—Rockwell Hardness Test—Part 3: Calibration of Reference Blocks (Scales A, B, C, D, E, F, G, H, K, N, T)
- ISO 3738-1, Hardmetals—Rockwell Hardness Test (Scale A)—Part 1: Test Method ISO 3738-2 Hardmetals—Rockwell Hardness Test (Scale A)—Part 2: Preparation and Calibration of Standard Test Blocks
- JIS B 7726, Rockwell Hardness Test—Verification of Testing Machines
- JIS B 7730, Rockwell Hardness Test—Calibration of Reference Blocks

CHAPTER 6

Tensile Testing

THE TENSILE TEST is one of the most commonly used tests for evaluating materials. In its simplest form, the tensile test is accomplished by gripping opposite ends of a test item within the load frame of a test machine. A tensile force is applied by the machine, resulting in the gradual elongation and eventual fracture of the test item. During this process, force-extension data, a quantitative measure of how the test item deforms under the applied tensile force, are usually monitored and recorded. The most important tensile test properties for the routine inspection of metals are: the yield strength; the ultimate tensile strength; and, a measure of ductility, either the percent elongation or the reduction in area.

Stress-Strain Behavior

During a tensile test, the force applied to the test piece and the length of elongation of the test piece are measured simultaneously. The applied force is measured by the test machine or by accessory force measuring devices. The amount of stretching or extension can be measured with an extensometer. An *extensometer* is a device used to measure the amount of stretch that occurs in a test piece. Because the amount of elastic stretch is quite small at or around the onset of yielding (in the order of 0.5% or less for steels), some manner of magnifying the stretch is required. An extensometer may be a mechanical device, in which case the magnification occurs by mechanical means. An extensometer may also be an electrical device, in which case the magnification may occur by mechanical means, electrical means, or by a combination of both. Extensometers generally have fixed gage lengths.

The applied force F and the extension ΔL are measured and recorded simultaneously at regular intervals, and the data pairs can be converted into a stress-strain diagram as shown in Fig. 1. The conversion from force-

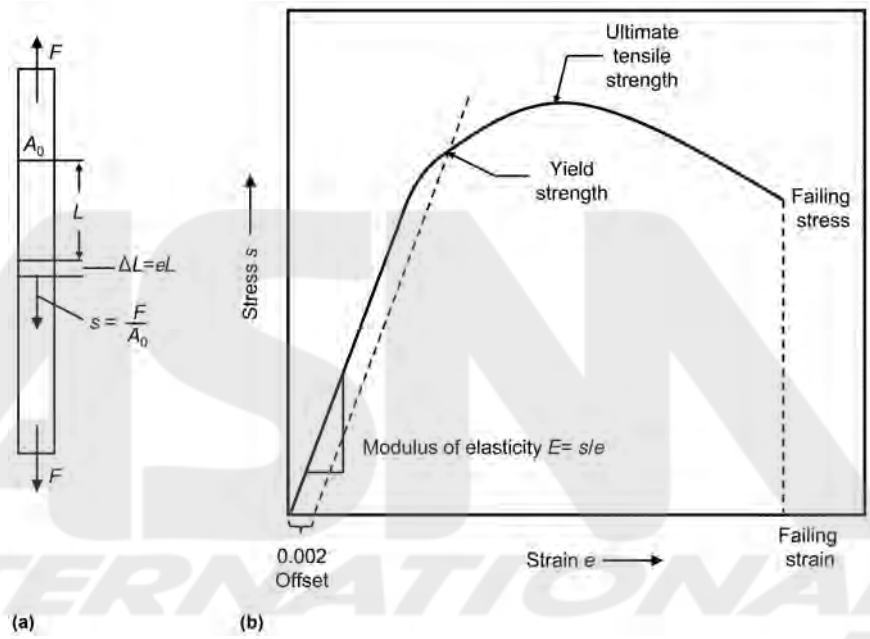


Fig. 1 Typical stress-strain behavior. (a) Definition of *s* and *e* in terms of initial test piece length *L*, and cross-sectional area A_0, before application of a tensile force *F*. (b) Stress-strain curve showing yield strength, ultimate tensile strength, and failing stress. Source: Adapted from Ref 1

extension data to stress-strain properties is shown schematically in Fig. 1(a). Engineering stress *s* is obtained by dividing the applied force by the original cross-sectional area A_0 of the test piece, and strain *e* is obtained by dividing the amount of extension ΔL by the original gage length *L*. The basic result is a stress-strain curve (Fig. 1b) with regions of elastic deformation and permanent or plastic deformation at stresses greater than those of the elastic limit.

Elastic deformation occurs in the initial portion of a stress-strain curve, where the stress-strain relationship is initially linear. In this region, the stress is proportional to strain. Mechanical behavior in this region of stress-strain curve is defined by the modulus of elasticity *E*. The modulus of elasticity is the slope of the stress-strain line in this linear region, and it is a basic physical property of all materials. Because the modulus of elasticity is a structure insensitive property that is not normally affected by processing, it is not normally reported on material certification sheets.

Properties from Test Results

Important properties determined from the tensile test include strength properties (yield strength and ultimate tensile strength) and, a measure of ductility, either percent elongation or reduction in area.

Strength Properties

Yield strength and tensile strength are the most common strength properties determined in a tensile test. According to ASTM E6, tensile strength is calculated from the maximum force during a tensile test that is carried to rupture divided by the original cross-sectional area of the test piece. The yield strength refers to the stress at which a small, but measurable, amount of inelastic or plastic deformation occurs. Yield strength is usually defined as:

- Upper yield strength or upper yield point
- Offset yield strength

An upper yield strength or upper yield point (Fig. 2a) usually occurs with low carbon steels and some other metal systems to a limited

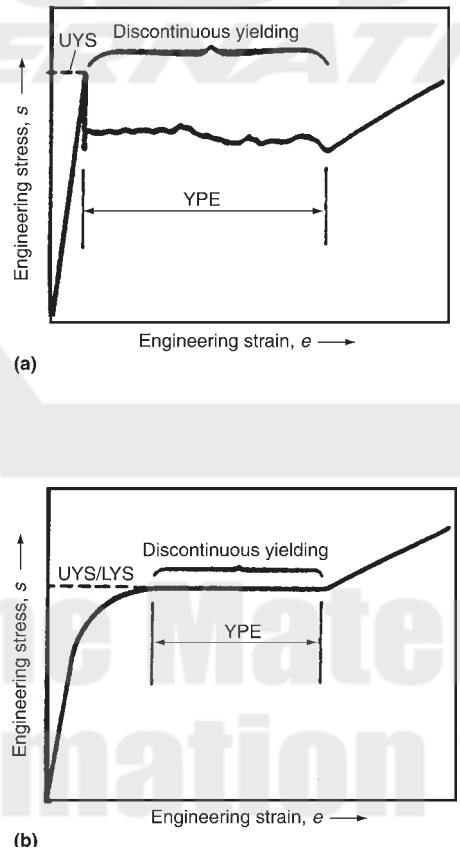


Fig. 2 Examples of stress-strain curves exhibiting pronounced yield point behavior. Pronounced yielding, of the type shown, is usually called yield point elongation (YPE). (a) Classic example of upper yield strength (UYS) behavior typically observed in low carbon steels with a very pronounced upper yield strength. (b) General example of pronounced yielding without an upper yield strength. LYS, lower yield strength. Source: Ref 2

degree. Often, the pronounced peak of the upper yield is suppressed due to slow testing speed or nonaxial loading (i.e., bending of the test piece), metallurgical factors, or a combination of these; in this case, a curve of the type shown in Fig. 2(b) is obtained. The offset definition of yield strength was developed for materials that do not exhibit the yield point behavior shown in Fig. 2. To determine the offset yield strength, the stress-strain curve must be determined during the test.

Upper yield strength or *upper yield point* can be defined as the stress at which measurable strain occurs without an increase in the stress; that is, there is a horizontal region of the stress-strain curve (Fig. 2) where discontinuous yielding occurs. Before the onset of discontinuous yielding, a peak of maximum stress for yielding is typically observed (Fig. 2a). This pronounced yielding, of the type shown, is usually called yield point elongation (YPE). This elongation is a diffusion related phenomenon, where under certain combinations of strain rate and temperature as the material deforms, interstitial atoms are dragged along with dislocations, or dislocations can alternately break away and be repinned, with little or no increase in stress. Either or both of these actions cause serrations or discontinuous changes in a stress-strain curve, which are usually limited to the onset of yielding. This type of yield point is sometimes referred to as the upper yield strength or upper yield point and is usually associated with low carbon steels, although other metal systems may exhibit yield points to some degree.

The yield point is easy to measure because the increase in strain that occurs without an increase in stress is visually apparent during the conduct of the test by observing the force indicating system. As shown in Fig. 2, the yield point is usually quite obvious and thus can easily be determined by observation during a tensile test. It can be determined from a stress-strain curve or by the halt of the dial when the test is performed on machines that use a dial to indicate the applied force. However, when watching the movement of the dial, sometimes a minimum value, recorded during discontinuous yielding, is noted. This value is sometimes referred to as the lower yield point. When the value is ascertained without instrumentation readouts, it is often referred to as the halt-of-dial or the drop-of-beam yield point (as an average usually results from eye readings). It is almost always the upper yield point that is determined from instrument readouts.

Offset yield strength is the stress that causes a specified amount of set to occur; that is, at this stress, the test piece exhibits plastic deformation or set equal to a specific amount. To determine the offset yield strength, it is necessary to secure data (autographic or numerical) from which a stress-strain diagram may be constructed graphically or in computer memory. Figure 3 shows how to use these data; the amount of the specified offset 0-*m* is laid out on the strain axis. A line, *m*-*n*, parallel to the modulus of elasticity line, 0-*A*, is drawn to intersect the stress-strain curve. The point

of intersection *r* is the offset yield strength, and the value *R* is read from the stress axis. Typically, for many materials, the offset specified is 0.2%; however, other values may be specified. Therefore, when reporting the offset yield strength, the amount of the offset also must be reported; for example, 0.2 % offset yield strength = 52.8 ksi or yield strength (0.2% offset) = 52.8 ksi, are common formats used in reporting this information.

Ductility

Ductility is the ability of a material to deform plastically without fracturing. A sketch of a test piece with a circular cross-section that has been pulled to fracture is shown in Fig. 4. As indicated in this sketch, the test

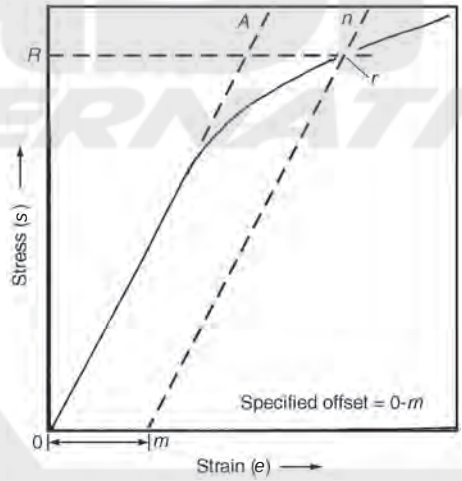


Fig. 3 Method of determining yield strength by the offset method (adaptation of Fig. 21 in ASTM E8). Source: Ref 2

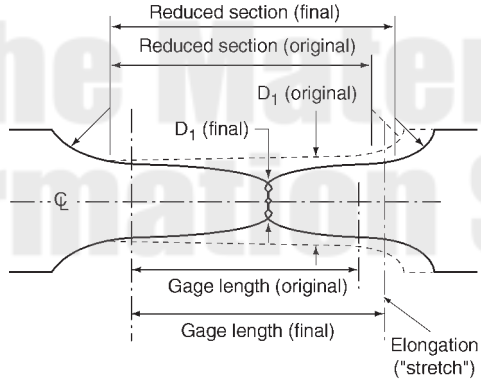


Fig. 4 Sketch of fractured, round tension test piece. Dashed lines show original shape. Strain = elongation/gage length. Source: Ref 2

piece elongates during the tensile test and correspondingly reduces in cross-sectional area. The two measures of the ductility of a material are the amount of elongation and reduction in area that occurs during a tensile test.

Elongation is defined in ASTM E6 as the increase in the gage length of a test piece subjected to a tension force, divided by the original gage length on the test piece. Elongation usually is expressed as a percentage of the original gage length. ASTM E6 further indicates the following:

- The increase in gage length may be determined either at or after fracture, as specified for the material under test
- The gage length shall be stated when reporting values of elongation
- Elongation is affected by test piece geometry (gage length, width, and thickness of the gage section and of adjacent regions) and test procedure variables, such as alignment and speed of pulling

The manual measurement of elongation on a tensile test piece can be done with the aid of gage marks applied to the unstrained reduced section. After the test, the amount of stretch between gage marks is measured with an appropriate device. The use of the term elongation in this instance refers to the total amount of stretch or extension. Elongation, in the sense of nominal engineering strain e is the value of gage extension divided by the original distance between the gage marks. Strain elongation is usually expressed as a percentage, where the nominal engineering strain is multiplied by 100 to obtain a percent value; that is:

$$e,\% = \left[\frac{(\text{final gage length} - \text{original gage length})}{\text{original gage length}}\right] \times 100$$

The final gage length at the completion of the test may be determined in two ways. Historically, it was determined manually by carefully fitting the two ends of the fractured test piece together and measuring the distance between the gage marks. However, some modern computer controlled testing systems obtain data from an extensometer that is left on the test piece through fracture. In this case, the computer may be programmed to report the elongation as the last strain value obtained prior to some event, perhaps the point at which the applied force drops to 90% of the maximum value recorded. There has been no general agreement about what event should be the trigger. Users and machine manufacturers find that different events may be appropriate for different materials, although some consensus has been reached, (see ASTM E8-99). The elongation values determined by these two methods are not the same. In general, the result obtained by the manual method is a couple of percent larger and is more variable because the test piece ends do not fit together perfectly. It is strongly recommended that when disagreements arise about elongation results, agreement should be reached on which method will be used prior to any further testing.

Effect of Gage Length and Necking. The effect of gage length on elongation values is shown in Fig. 5. Gage length is very important; however, as the gage length becomes quite large, the elongation tends to be independent of the gage length. The gage length must be specified prior to the test, and it must be shown in the data record for the test.

Figures 4 and 5 also illustrate considerable localized deformation in the vicinity of the fracture. This region of local deformation is often called a neck, and the occurrence of this deformation is termed necking. Necking occurs as the force begins to drop after the maximum force has been reached on the stress-strain curve. Up to the point at which the maximum force occurs, the strain is uniform along the gage length; that is, the strain is independent of the gage length. However, once necking begins, the gage length becomes very important. When the gage length is short, this localized deformation becomes the principal portion of measured elongation. For long gage lengths, the localized deformation is a much smaller portion of the total. For this reason, when elongation values are reported, the gage length must also be reported, for example, elongation = 25% (50 mm, or 2.00 in., gage length).

Reduction of area is another measure of the ductility of metal. As a test piece is stretched, the cross-sectional area decreases, and as long as

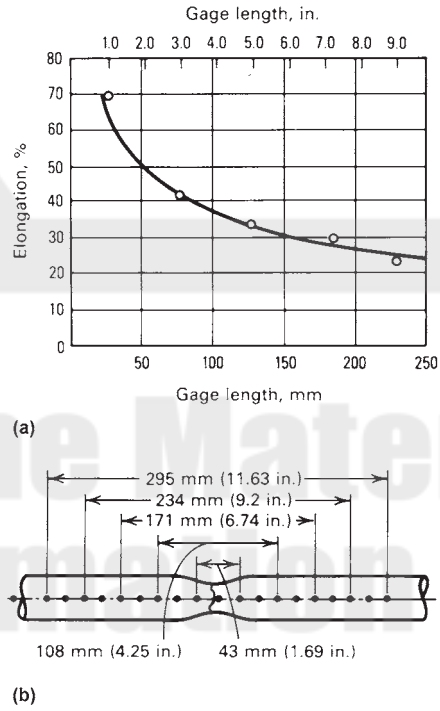


Fig. 5 Effect of gage length on the percent elongation. (a) Elongation, %, as a function of gage length for a fractured tension test piece. (b) Distribution of elongation along a fractured tension test piece. Original spacing between gage marks, 12.5 mm (0.5 in.). Source: Ref 3

the stretch is uniform, the reduction of area is proportional to the amount of stretch or extension. However, once necking begins to occur, proportionality is no longer valid.

According to ASTM E6, *reduction of area* is defined as "the difference between the original cross-sectional area of a tensile test piece and the area of its smallest cross section." Reduction of area is usually expressed as a percentage of the original cross-sectional area of the test piece. The smallest final cross section may be measured at or after fracture as specified for the material under test. The reduction of area (RA) is almost always expressed as a percentage:

$$\text{RA},\% = \left[\frac{(\text{original area} - \text{final area})}{\text{original area}}\right] \times 100$$

Reduction of area is customarily measured only on test pieces with an initial circular cross section because the shape of the reduced area remains circular or nearly circular throughout the test for such test pieces. In contrast, with rectangular test pieces, the corners prevent uniform flow from occurring, and consequently, after fracture, the shape of the reduced area is not rectangular (Fig. 6). Although a number of expressions have been used in an attempt to describe the way to determine the reduced area, none has received general agreement. Thus, if a test specification requires the measurement of the reduction of area of a test piece that is not circular, the method of determining the reduced area should be agreed to prior to performing the test.

Testing Machines

Conventional test machines for measuring mechanical properties include tensile testers, compression testers, or the more versatile universal testing machine (UTM). UTMs have the capability to test material in tension, compression, or bending. The word *universal* refers to the variety of stress states that can be studied. Universal testing machines can load material with a single, continuous or monotonic pulse, or in a cyclic manner. Other conventional test machines may be limited to either tensile loading or compressive loading, but not both. These machines have less versatility than UTM equipment, but are less expensive to purchase and maintain.

Fig. 6 Sketch of end view of rectangular test piece after fracture showing constraint at corners indicating the difficulty of determining reduced area. Source: Ref 2

The basic aspects of UTM equipment and testing generally apply to tension or compression testing machines as well.

Although there are many types of test systems in current use, the most common are universal testing machines, which are designed to test specimens in tension, compression, or bending. The testing machines are designed to apply a force to a material to determine its strength and resistance to deformation. Regardless of the method of force application, testing machines are designed to drive a crosshead or platen at a controlled rate, thus applying a tensile or compressive load to a specimen. Such testing machines measure and indicate the applied force in pound-force (lbf), kilogram-force (kgf), or newtons (N).

The load applying mechanism may be a hydraulic piston and cylinder with an associated hydraulic power supply or the load may be administered via precision cut machine screws driven by the necessary gears, reducers, and motor to provide a suitable travel speed. In some light capacity machines (only a few hundred pounds maximum), the force is applied by an air piston and cylinder. Gear driven systems obtain load capacities up to approximately 600 kN (1.35×105 lbf), while hydraulic systems can obtain forces up to approximately 4500 kN (1×106 lbf).

Conventional gear driven systems are generally designed for speeds of about 0.001 to 500 mm/min (4×10^{-6} to 20 in./min), which is suitable for quasi-static testing. Servohydraulic systems are generally designed over a wider range of test speeds.

Gear driven (or screw driven) machines are electromechanical devices that use a large actuator screw threaded through a moving crosshead (Fig. 7). The screw is turned in either direction by an electric motor through a gear reduction system. The screws are rotated by a variable control motor and drive the moveable crosshead up or down. This motion can load the specimen in either tension or compression, depending on how the specimen is to be held and tested.

A range of crosshead speeds can be achieved by varying the speed of the electric motor and by changing the gear ratio. A closed loop servodrive system ensures that the crosshead moves at a constant speed. The desired or user selected speed and direction information is compared with a known reference signal, and the servomechanism provides positional control of the moving crosshead to reduce any error or difference. State-of-the-art systems use precision optical encoders mounted directly on preloaded twin ball screws. These types of systems are capable of measuring crosshead displacement to an accuracy of 0.125% or better with a resolution of 0.6 μm.

Servohydraulic machines use a hydraulic pump and servohydraulic valves that move an actuator piston (Fig. 8). The actuator piston is attached to one end of the specimen. The motion of the actuator piston can be controlled in both directions to conduct tension, compression, or cyclic loading tests.

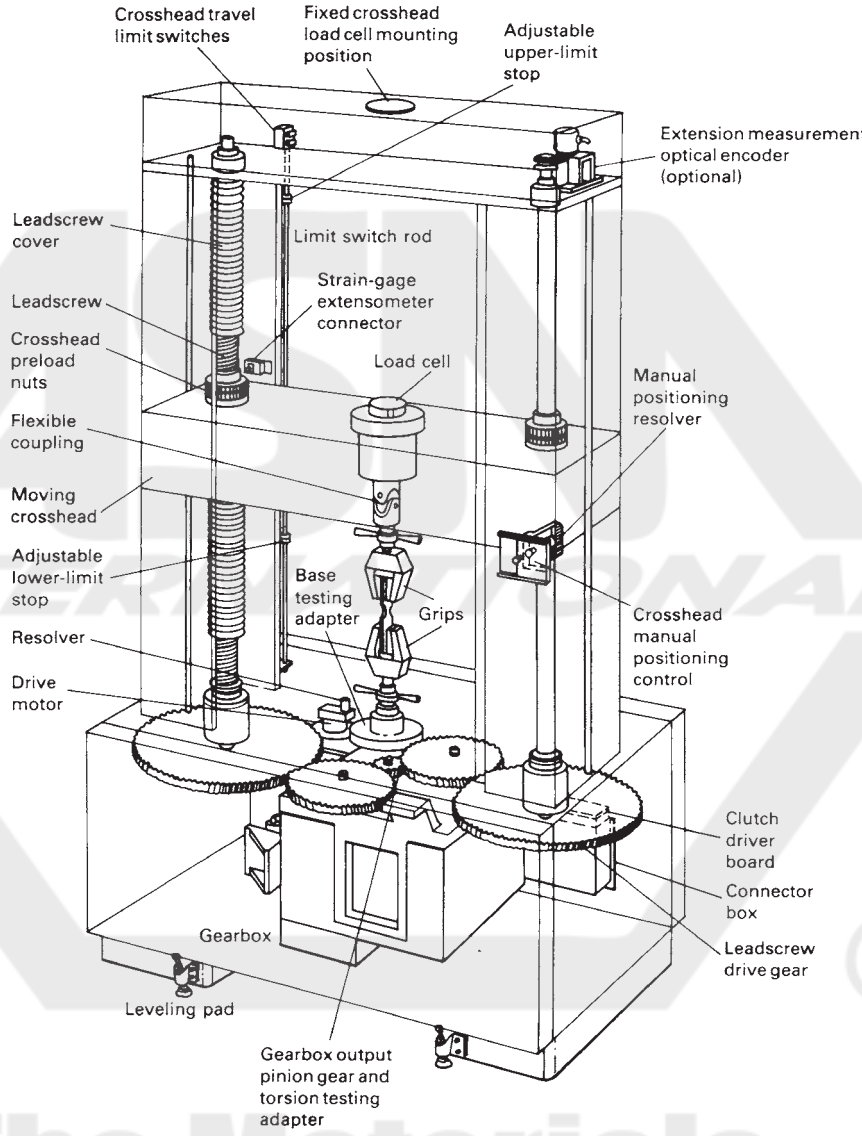


Fig. 7 Components of an electromechanical (screw driven) testing machine. For the configuration shown, moving the lower (intermediate) head upward produces tension in the lower space between the crosshead and the base. Source: Ref 4

Servohydraulic test systems have the capability of testing at rates from as low as 45×1011 m/s (1.8×10^{-9} in./s) to 30 m/s (1200 in./s) or more. The actual useful rate for any particular system depends on the size of the actuator, the flow rating of the servovalve, and the noise level present in the system electronics. A typical servohydraulic UTM system is shown in Fig. 9.

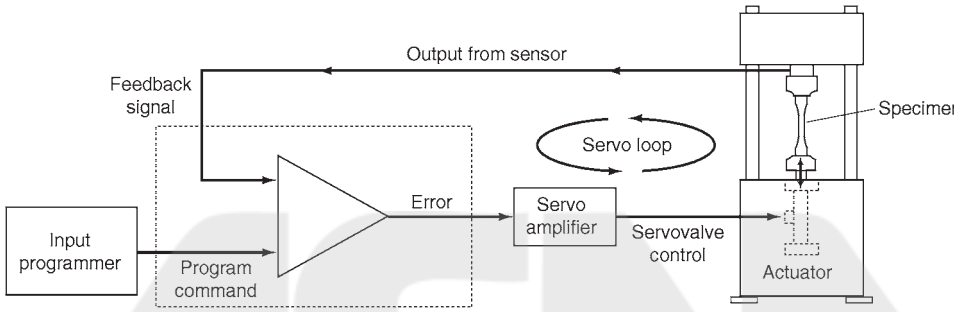


Fig. 8 Schematic of a basic servohydraulic, closed loop testing machine. Source: Ref 4

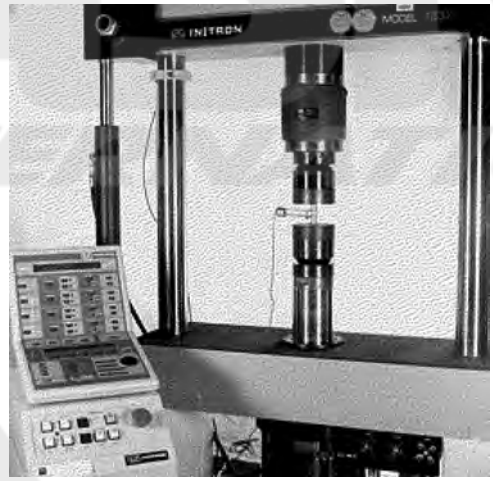


Fig. 9 Servohydraulic testing machine and load frame with a dedicated microprocessor-based controller. Source: Ref 4

Hydraulic actuators are available in a wide variety of force ranges. They are unique in their ability to economically provide forces of 4450 kN (1,000,000 lbf) or more. Screw driven machines are limited in their ability to provide high forces due to problems associated with low machine stiffness and large and expensive loading screws, which are increasingly more difficult to produce as the force rating goes up.

Microprocessors for Testing and Data Reduction. Contemporary UTMs are controlled by microprocessor-based electronics. One class of controllers is based on dedicated microprocessors for test machines (Fig. 9). Dedicated microprocessors are designed to perform specific tasks and have displays and input functions that are limited to those tasks. The dedicated microprocessor sends signals to the experimental apparatus and receives information from various sensors. The data received from sensors can be passed to oscilloscopes or computers for display and storage.

The experimental results consist of time and voltage information that must be further reduced to analyze material behavior. Analysis of the data requires the conversion of test results, such as voltage, to specific quantities, such as displacement and load, based on known conversion factors.

The second class of controllers is the personal computer (PC) designed with an electronic interface to the experimental apparatus, and the appropriate application software. The software takes the description of the test to be performed, including specimen geometry data, and establishes the requisite electronic signals. Once the test is underway, the computer controls the tests and collects, reduces, displays, and stores the data. The obvious advantage of the PC-based controller is reduced time to generate graphic results, or reports. The other advantage is the elimination of some procedural errors, or the reduction of the interfacing details between the operator and the experimental apparatus. Some systems are designed with both types of controllers. Having both types of controllers provides maximum flexibility in data gathering with a minimal amount of time required for reducing data when conducting standard experiments.

Measuring Load. Current testing machines use strain gaged load cells and pressure transducers. In a load cell, strain gages are mounted on precision machined alloy steel elements, hermetically sealed in a case with the necessary electrical outlets, and arranged for tensile and/or compressive loading. The load cell can be mounted so that the specimen is in direct contact, or the cell can be indirectly loaded through the machine crosshead, table, or columns of the load frame. The load cell and the load cell circuit are calibrated to provide a specific voltage as an output signal when a certain force is detected. In pressure transducers, which are variations of strain gage load cells, the strain gaged member is activated by the hydraulic pressure of the system.

Strain Measurement. Deformation of the specimen can be measured in several ways, depending on the size of specimen, environmental conditions, and measurement requirements for accuracy and precision of anticipated strain levels. A simple method is to use the velocity of the crosshead while tracking the load as a function of time. For the load and time data pair, the stress in the specimen and the amount of deformation, or strain, can be calculated. When the displacement of the platen is assumed to be equal to the specimen displacement, an error is introduced by the fact that the entire load frame has been deflected under the stress state. This effect is related to the machine stiffness.

The elongation of a specimen during load application can be measured directly with various types of devices, such as clip-on extensometers, directly mounted strain gages, and various optical devices. These devices are used extensively and can provide a high degree of deformation (strain) measurement accuracy. Other more advanced instrumentations, such as laser interferometry and video extensometers, are also available.

General Procedures

Numerous groups have developed standard methods for conducting the tensile test. In the United States, standards published by ASTM are commonly used to define tensile test procedures and parameters. Of the various ASTM standards related to tensile tests (see "Selected References" listed at the end of this chapter), the most common method for tensile testing of metallic materials is ASTM E8 "Standard Test Methods for Tension Testing of Metallic Materials" or the version using metric units, ASTM E8M. Standard methods for conducting the tensile test are also available from other standards organizations, such as the Japanese Industrial Standards (JIS), the Deutsche Institut für Normung (DIN), and the International Organization for Standardization (ISO). Other domestic technical groups in the United States have developed standards, but in general, these are based on ASTM E8.

With the increasing internationalization of trade, methods developed by other national standards organizations (such as JIS, DIN, or ISO standards) are increasingly being used in the United States. Although most tension test standards address the same concerns, they differ in the values assigned to variables. Thus, a tension test performed in accordance with ASTM E8 will not necessarily have been conducted in accordance with ISO 6892 or JIS Z2241, and so on, and vice versa. Therefore, it is necessary to specify the applicable testing standard for any test results or mechanical property data.

Unless specifically indicated otherwise, the values of all variables discussed hereafter are those related to ASTM E8 "Standard Test Methods for Tension Testing of Metallic Materials." The test consists of three distinct parts:

1. Test piece preparation, geometry, and material condition
2. Test setup and equipment
3. Test

The Test Piece

The test piece is one of two basic types. Either it is a full cross section of the product form, or it is a small portion that has been machined to specific dimensions. Full section test pieces consist of a part of the test unit as it is fabricated. Examples of full section test pieces include bars, wires, and hot rolled or extruded angles cut to a suitable length and then gripped at the ends and tested. In contrast, a machined test piece is a representative sample, such as one of the following:

- Test piece machined from a rough specimen taken from a coil or plate
- Test piece machined from a bar with dimensions that preclude testing a full section test piece because a full section test piece exceeds the

capacity of the grips or the force capacity of the available testing machine or both
- Test piece machined from material of great monetary or technical value

In these cases, representative samples of the material must be obtained for testing. The descriptions of the tensile test proceed from the point that a rough specimen (Fig. 10) has been obtained. That is, the rough specimen has been selected based on some criteria, usually a material specification or a test order issued for a specific reason.

In this chapter, the term test piece is used for what is often called a specimen. This terminology is based on the convention established by ISO Technical Committee 17, Steel in ISO 377-1, "Selection and Preparation of Samples and Test Pieces of Wrought Steel," where terms for a *test unit*, a *sample product*, *sample*, *rough specimen*, and *test piece* are defined as follows:

- *Test unit*: The quantity specified in an order that requires testing (for example, 10 tons of in. bars in random lengths)
- *Sample product*: Item (in the previous example, a single bar) selected from a test unit for the purpose of obtaining the test pieces
- *Sample*: A sufficient quantity of material taken from the sample product for the purpose of producing one or more test pieces. In some cases, the sample may be the sample product itself (i.e., a 2 ft length of the sample product).
- *Rough specimen*: Part of the sample having undergone mechanical treatment, followed by heat treatment where appropriate, for the purpose of producing test pieces; in the example, the sample is the rough specimen.
- *Test piece*: Part of the sample or rough specimen, with specified dimensions, machined or unmachined, brought to the required condition for submission to a given test. If a testing machine with sufficient force capacity is available, the test piece may be the rough specimen; if sufficient capacity is not available, or for another reason, the test piece may be machined from the rough specimen to dimensions specified by a standard.

These terms are shown graphically in Fig. 10. As shown, the test piece, or what is commonly called a specimen, is a very small part of the entire test unit.

Description of Test Material

Test Piece Orientation. Orientation and location of a test material from a product can influence measured tensile properties. Although modern metal working practices, such as cross rolling, have tended to reduce

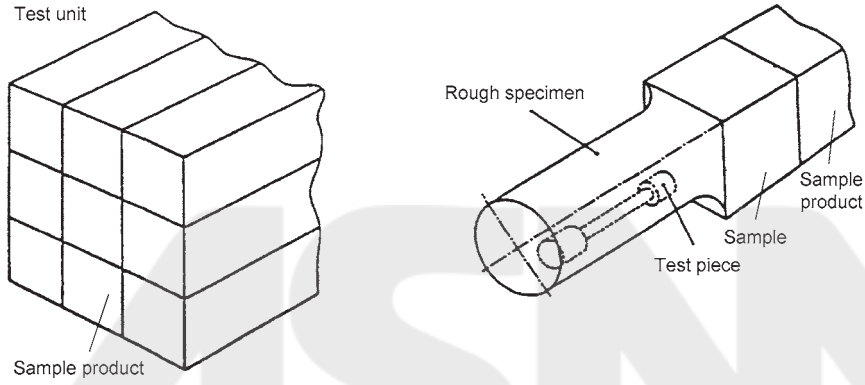


Fig. 10 Illustration of ISO terminology used to differentiate between sample, specimen, and test piece (see text for definitions of test unit, sample product, sample, rough specimen, and test piece). As an example, a test unit may be a 250 ton heat of steel that has been rolled into a single thickness of plate. The sample product is thus one plate from which a single test piece is obtained. Source: Ref 2

the magnitude of the variations in the tensile properties, it must not be neglected when locating the test piece within the specimen or the sample.

Because most materials are not isotropic, test piece orientation is defined with respect to a set of axes as shown in Fig. 11. These terms for the orientation of the test-piece axes in Fig. 11 are based on the convention used by ASTM E8 "Fatigue and Fracture." This scheme is identical to that used by the ISO Technical Committee 164 "Mechanical Testing," although the L, T, and S axes are referred to as the X, Y, and Z axes, respectively, in the ISO documents.

When a test is being performed to determine conformance to a product standard, the product standard must state the proper orientation of the test piece with regard to the axis of prior working, (e.g., the rolling direction of a flat product). Because alloy systems behave differently, no general rule of thumb can be stated on how prior working may affect the directionality of properties. As can be seen in Table 1, the longitudinal strengths of steel are generally somewhat less than the transverse strength. However, for aluminum alloys, the opposite is generally true.

Many standards, such as ASTM A370, E8, and B557, provide guidance in the selection of test piece orientation relative to the rolling direction of the plate or the major forming axes of other types of products and in the selection of specimen and test-piece location relative to the surface of the product. Orientation is also important when characterizing the directionality of properties that often develops in the microstructure of materials during processing. For example, some causes of directionality include the fibering of inclusions in steels, the formation of crystallographic textures in most metals and alloys, and the alignment of molecular chains in polymers.

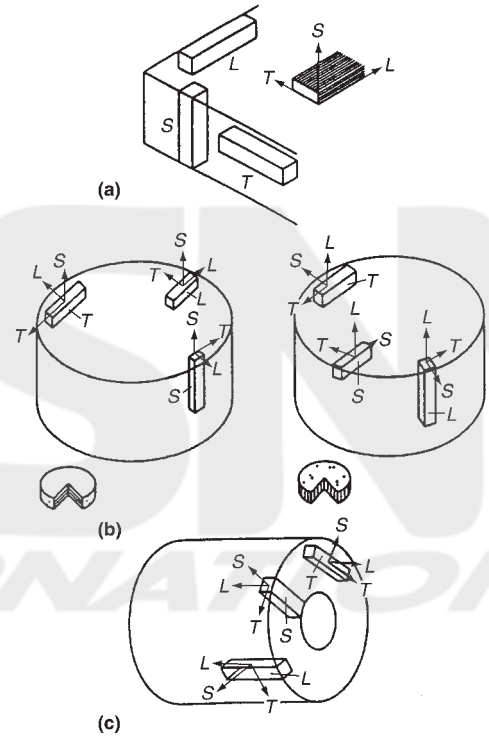


Fig. 11 System for identifying the axes of test-piece orientation in various product forms. (a) Flat rolled products. (b) Cylindrical sections. (c) Tubular products. Source: Ref 2

Table 1 Effect of test piece orientation on tensile properties

Orientation	Yield strength, ksi	Tensile strength, ksi	Elongation in 50 mm (2 in.), %	Reduction of area, %
ASTM A 572, Grade 50 (¾ in. thick plate, low sulfur level)				
Longitudinal	58.8	84.0	27.0	70.2
Transverse	59.8	85.2	28.0	69.0
ASTM A 656, Grade 80 (¾ in. thick plate, low sulfur level + controlled rolled)				
Longitudinal	81.0	102.3	25.8	71.2
Transverse	86.9	107.9	24.5	67.1
ASTM A 5414 (¾ in. thick plate, low sulfur level)				
Longitudinal	114.6	121.1	19.8	70.6
Transverse	116.3	122.2	19.5	69.9

Source: Courtesy of Francis J. Marsh. Source: Ref 2

The location from which a test material is taken from the initial product form is important because the manner in which a material is processed influences the uniformity of microstructure along the length of the product as well as through its thickness properties. For example, the properties of metal cut from castings are influenced by the rate of cooling and by shrink-

age stresses at changes in the section. Generally, test pieces taken from near the surface of iron castings are stronger. To standardize test results relative to location, ASTM A370 recommends that tensile test pieces be taken from midway between the surface and the center of round, square, hexagon, or octagonal bars. ASTM E8 recommends that test pieces be taken from the thickest part of a forging from which a test coupon can be obtained, from a prolongation of the forging, or in some cases, from separately forged coupons representative of the forging.

Test Piece Geometry

As previously noted, the item being tested may be either the full cross section of the item or a portion of the item that has been machined to specific dimensions. Test piece geometry is often influenced by product form. For example, only test pieces with rectangular cross sections can be obtained from sheet products. Test pieces taken from thick plate may have either flat (plate type) or round cross sections. Most tensile test specifications show machined test pieces with either circular cross sections or rectangular cross sections. Nomenclature for the various sections of a machined test piece is shown in Fig. 12. Most tensile test specifications present a set of dimensions, for each cross-section type, that are standard, as well as additional sets of dimensions for alternative test pieces. In general, the standard dimensions published by ASTM, ISO, JIS, and DIN are similar, but they are not identical.

Measurement of Initial Test Piece Dimensions. Machined test pieces are expected to meet size specifications, but to ensure dimensional accuracy, each test piece should be measured prior to testing. Gage length, fillet radius, and cross-sectional dimensions are measured easily. Cylindrical test pieces should be measured for concentricity. Maintaining acceptable concentricity is extremely important in minimizing unintended bending stresses on materials in a brittle state.

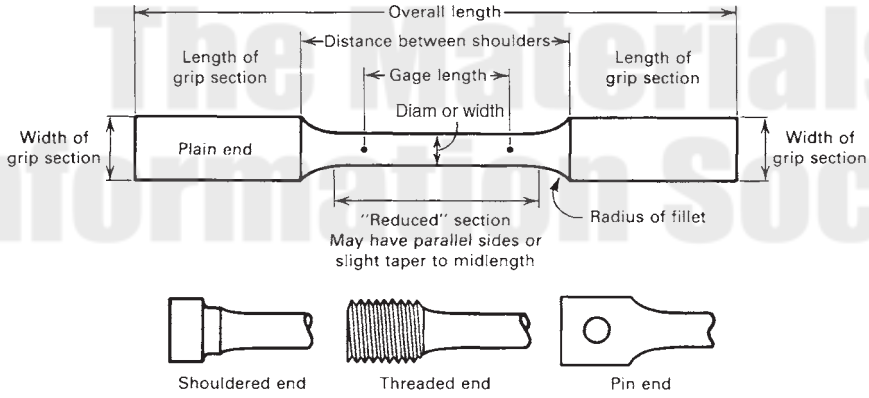


Fig. 12 Nomenclature for a typical tension test piece. Source: Ref 2

Measurement of Cross-Sectional Dimensions. The test pieces must be measured to determine whether they meet the requirements of the test method. Test piece measurements must also determine the initial cross-sectional area when it is compared against the final cross section after testing as a measure of ductility. The precision with which these measurements are made is based on the requirements of the test method, or if none are given, on good engineering judgment.

Measurement of the Initial Gage Length. ASTM E8 assumes that the initial gage length is within specified tolerance; therefore, it is necessary only to verify that the gage length of the test piece is within the tolerance.

Marking Gage Length. Measurement of elongation requires marking the gage length of the test piece. The gage marks should be placed on the test piece in a manner so that when fracture occurs, the fracture will be located within the center one-third of the gage length (or within the center one-third of one of several sets of gage length marks). For a test piece machined with a reduced section length that is the minimum specified by ASTM E8 and with a gage length equal to the maximum allowed for that geometry, a single set of marks is usually sufficient.

Surface Finish and Condition. Test pieces from materials that are not high strength or that are ductile are usually insensitive to surface finish effects. However, if surface finish in the gage length of a tensile test piece is extremely poor (with machine tool marks deep enough to act as stress-concentrating notches, for example), test results may exhibit a tendency toward decreased and variable strength and ductility.

It is good practice to examine the test piece surface for deep scratches, gouges, edge tears, or shear burrs. These discontinuities may sometimes be minimized or removed by polishing or, if necessary, by further machining; however, dimensional requirements may often no longer be met after additional machining or polishing. In all cases, the reduced sections of machined test pieces must be free of detrimental characteristics, such as cold work, chatter marks, grooves, gouges, burrs, and so on.

Test Setup

The setup of a tensile test involves the installation of a test piece in the load frame of suitable factors; for example, calibration and load frame rigidity must be considered. The other aspects of the test setup include proper gripping and alignment of the test piece as well as the installation of extensometers or strain sensors when plastic deformation or yield behavior of the piece is being measured.

Gripping Devices. The grips must furnish an axial connection between the test piece and the testing machine; that is, the grips must not cause bending in the test piece during loading. The choice of grip is primarily

dependent on the geometry of the test piece and, to a lesser degree, on the preference of the test laboratory. The load capacities of grips range from under 4.5 kgf (10 lbf) to 45,000 kgf (100,000 lbf), or more. ASTM E8 describes the various types of gripping devices used to transmit the measured load applied by the test machine to the tensile test specimen. Several of the many grips that are in common use are illustrated in Fig. 13, but many other designs are also used. As shown, the gripping devices can be classified into several distinct types, wedges, threaded, button, and snubbing.

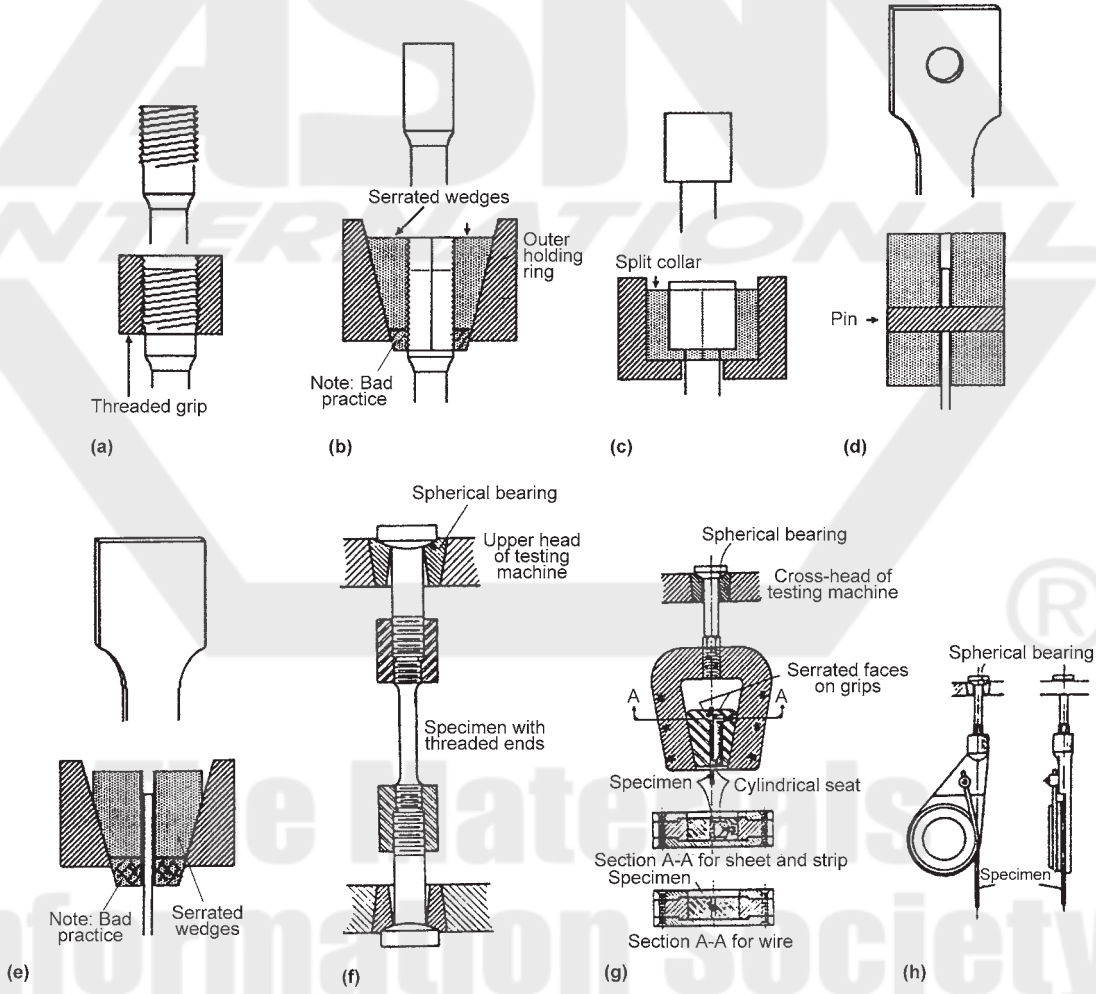


Fig. 13 Examples of gripping methods for tension test pieces. (a) Round specimen with threaded grips. (b) Gripping with serrated wedges with hatched region showing bad practice of wedges extending below the outer holding ring. (c) Butt end specimen constrained by a split collar. (d) Sheet specimen with pin constraints. (e) Sheet specimen with serrated wedge grip with hatched region showing the bad practice of wedges extended below the outer holding ring. (f) Gripping device for threaded end specimen. (g) Gripping device for sheet and wire. (h) Snubbing device for testing wire. Sources: Adapted from Ref 5 and ASTM E8

Alignment of the Test Piece. The force application axis of the gripping device must coincide with the longitudinal axis of symmetry of the test piece. If these axes do not coincide, the test piece will be subjected to a combination of axial loading and bending. For ductile materials, the effect of bending is minimal, other than the suppression of the upper yield stress. However, if the material has little ductility, the increased strain due to bending may cause fracture to occur at a lower stress than if there were no bending.

Extensometers. When the tension test requires the measurement of strain behavior (i.e., the amount of elastic and/or plastic deformation occurring during loading), extensometers or strain gages must be attached to the test piece. The amount of strain can be quite small (e.g., approximately 0.5% or less for elastic strain in steels), and extensometers and other strain sensing systems, such as strain gages, are designed to magnify strain measurement into a meaningful signal for data processing.

The elongation of a specimen during load application can be measured directly with various types of devices, such as clip-on extensometers (Fig. 14), directly mounted strain gages (Fig. 15), and various optical devices. Several types of extensometers are available. Extensometers generally have fixed gage lengths. If an extensometer is used only to obtain a portion of the stress-strain curve sufficient to determine the yield properties, the gage length of the extensometer may be shorter than the gage length required for the elongation-at-fracture measurement. It may also be longer, but in general, the extensometer gage length should not exceed approximately 85% of the length of the reduced section or the distance between the grips for test pieces without reduced sections.

Fig. 14 Test specimen with an extensometer attached to measure specimen elongation. Courtesy of Epsilon Technology Corporation. Source: Ref 4

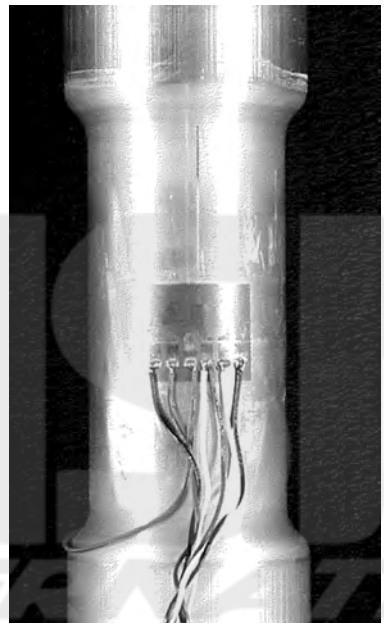


Fig. 15 Strain gages mounted directly to a specimen. Source: Ref 4

Test Procedures

After the test piece has been properly prepared and measured and the test setup established, conducting the test is fairly routine. The test piece is installed properly in the grips, and if required, extensometers or other strain measuring devices are fastened to the test piece for measurement and recording of extension data. Data acquisition systems also should be checked. In addition, it is sometimes useful to repetitively apply small initial loads and vibrate the load train (a metallographic engraving tool is a suitable vibrator) to overcome friction in various couplings. A check can also be run to ensure that the test will run at the proper testing speed and temperature. The test is then begun by initiating force application.

Post Test Measurements

After the test has been completed, it is often required that the cross-sectional dimensions again be measured to obtain measures of ductility. ASTM E8 states that measurements made after the test shall be to the same accuracy as the initial measurements.

Method E8 also states that upon completion of the test, gage lengths 2 in. and under are to be measured to the nearest 0.01 in., and gage lengths over 2 in. are to be measured to the nearest 0.5%. The document goes on to state that a percentage scale reading to 0.5 % of the gage length may be used. However, if the tension test is being performed as part of a product

specification, and the elongation is specified to meet a value of 3% or less, special techniques, which are described, are to be used to measure the final gage length.

ACKNOWLEDGMENT

This chapter was adapted from Uniaxial Tension Testing by J.M. Holt in *Mechanical Testing and Evaluation*, Volume 8, *ASM Handbook*, 2000.

REFERENCES

1. F.C. Campbell, *Elements of Metallurgy and Engineering Alloys*, ASM International, 2008
2. J. M. Holt, Uniaxial Tension Testing, *Mechanical Testing and Evaluation*, Vol 8, *ASM Handbook*, ASM International, 2000
3. *Making, Shaping, and Treating of Steel*, 10th ed., U.S. Steel, 1985, Fig. 50-12 and 50-13
4. J.W. House and P.P. Gillis, Testing Machines and Strain Sensors, *Mechanical Testing and Evaluation*, Vol 8, *ASM Handbook*, ASM International, 2000
5. D. Lewis, Tensile Testing of Ceramics and Ceramic-Matrix Composites, *Tensile Testing*, P. Han, Ed., ASM International, 1992, p 147–182

SELECTED REFERENCES

- "Standard Test Methods for Tension Testing of Metallic Materials." E8, ASTM
- "Standard Methods and Definitions for Mechanical Testing of Steel Products," A370, ASTM
- "Standard Test Methods for Poisson's Ratio at Room Temperature," E132, ASTM
- "Standard Test Methods for Young's Modulus, Tangent Modulus, and Chord Modulus," E111, ASTM
- "Standard Methods of Tension Testing of Metallic Materials," E8, ASTM
- "Standard Methods of Tension Testing Wrought and Cast Aluminum- and Magnesium-Alloy Products," B557, ASTM
- "Standard Recommended Practice for Verification of Specimen Alignment Under Tensile Loading," E1012, ASTM

CHAPTER 7

Chemical Composition

THE OVERALL CHEMICAL COMPOSITION of metals and alloys is most commonly determined by x-ray fluorescence (XRF) and optical emission spectroscopy (OES). While these methods work well for most elements, they are not useful for dissolved gases and some nonmetallic elements that can be present in metals as alloying or impurity elements. High-temperature combustion and inert gas fusion methods are typically used to analyze dissolved gases (oxygen, nitrogen, hydrogen) and in some cases, carbon and sulfur in metals.

X-Ray Fluorescence Spectroscopy (XRF)

X-ray fluorescence is capable of the detection and quantification of elements with atomic number 5 or higher. Older energy dispersive units with beryllium window detectors are limited to atomic number 11 or higher.

Typical uses are:

- Qualitative and quantitative chemical analysis for major and minor elements in metals and alloys
- Determination of composition and thickness of thin film deposits

X-ray fluorescence with energy dispersive detectors have a threshold sensitivity of ~0.02% and a precision for quantitative analysis of ~1% relative, or 0.02% absolute, depending on count time. The detection threshold and precision for wavelength dispersive detectors is a threshold sensitivity of ~0.005% and a precision for quantitative analysis of ~0.2% relative, or 0.005% absolute, whichever is greater.

For powders, the sample size is several grams pressed into a pellet. Powder samples are typically attached to substrates not produced by x-ray or are pressed into pellets. For bulk solids, the sample size is typically ~1 cm (~0.4 in.) diameter spot on the surface with a depth of 10 to 100 μm,

which increases with decreasing average atomic weight of the sample. Bulk metal samples typically are ground to produce a flat uncontaminated surface for analysis. Typical samples have dimensions of several centimeters; however, most instruments can accommodate samples 10 cm (4 in.) or more in diameter. Some instruments are designed to map compositional variations, such as segregation in cross-sectioned ingots. These instruments have variable beam diameters down to ~0.1 mm (~0.004 in.), and *x-y* stages to translate the sample under the beam.

The two main limitations are:

- Elements with low atomic numbers produce very few x-rays and are difficult to detect or quantify, particularly in energy dispersive systems. Wavelength dispersive instruments can go down to atomic number 5 (boron). Modern energy dispersive systems are limited to atomic number ~7 (nitrogen) and above. Older energy dispersive systems cannot readily analyze for elements with atomic number less than 11 (sodium).
- Some combinations of elements are difficult to analyze because of overlapping x-ray energies. Such problems sometimes can be overcome by using wavelength dispersive spectrometers, rather than energy dispersive detectors, or by using optical emission spectroscopy.

Operating Principles

Physical Basis. The negatively charged electrons surrounding each atom's nucleus exist in discrete energy levels or orbitals, as shown in Fig. 1.

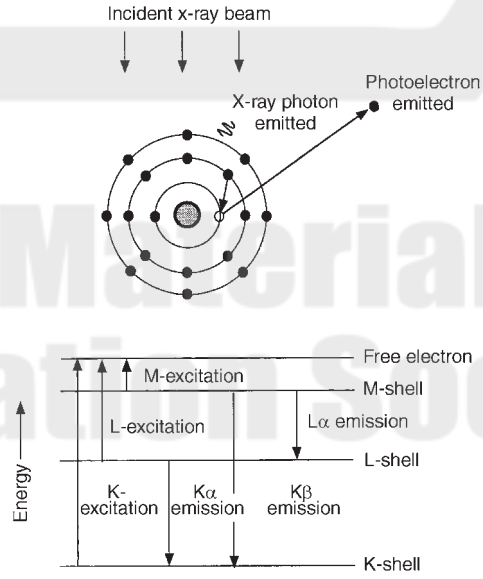


Fig. 1 Excitation of atom in sample by ejection of electrons or photoelectron production, and relaxation of excited atom by electronic transitions and accompanying characteristic x-ray emission. Source: Ref 1

Each electron's energy depends on its quantum state or the orbital it inhabits, and the number of positively charged protons in the atom's nucleus. Electrons with low principal quantum numbers (those close to the nucleus) are tightly bound; they require large amounts of energy to remove them (i.e., to cause ionization). Electrons with higher principal quantum numbers (those farther from the nucleus) are less tightly bound; less energy is required to remove them. Atoms with many positive protons in the nucleus (elements of high atomic number) tend to bind their inner shell electrons more tightly. Atoms with few protons in the nucleus (elements of lower atomic number) are more easily stripped of their inner shell electrons.

The net result is that each element has a unique set of known electron energy levels. Similarly, the set of energy differences between these electron energy levels are also unique for each element and constitutes a characteristic *fingerprint* by which each element can be identified. In XRF spectroscopy, as well as many other analytical methods, the combined electron energy level *fingerprints* of the elements present in the sample are experimentally obtained and are then compared to the fingerprints of known elements. From these comparisons, it is possible to identify the elements and their compositions present in a sample.

When sufficient energy is externally supplied to a sample of unknown composition, some of the electrons are excited to higher quantum states or energy levels, or removed from the atom or ionization. These *excited* atoms quickly *relax* by electrons from higher energy levels filling the vacated levels. When this happens, photons are emitted whose energies are equal to the differences between the two energy levels involved; this process is called *fluorescence*. If the energies of these emitted photons are measured, they provide the *fingerprint* of the unknown sample. This measurement can then be compared to the known *fingerprints* of the elements, enabling determination of which elements are present in the sample and the concentration of each element present. Which elements are present can be deduced from the energies of the photons emitted by the sample. How much of each element is present can be deduced from the numbers of photons with energies characteristic of the various elements. The characteristic energies emitted as excited atoms relax span a range of the electromagnetic spectrum. Electronic transitions between inner shells typically produce x-rays (photons with energies in the 200 to 20,000 eV range, characterized by wavelengths of 6 to 0.06 nm).

Instrumentation. A simplified schematic of an XRF spectrometer is shown in Fig. 2. A beam of x-rays is produced by electronic excitation of a metal target in the instrument's x-ray source. The beam's only function is to excite atoms in the sample. As the incident x-ray beam is directed onto the sample surface, it penetrates some small distance into the sample, typically 10 to 100 μm, depending on the atomic numbers of the elements in the sample. Penetration depths are greater for low atomic number elements. The incident x-rays excite atoms in the sample, transitioning some

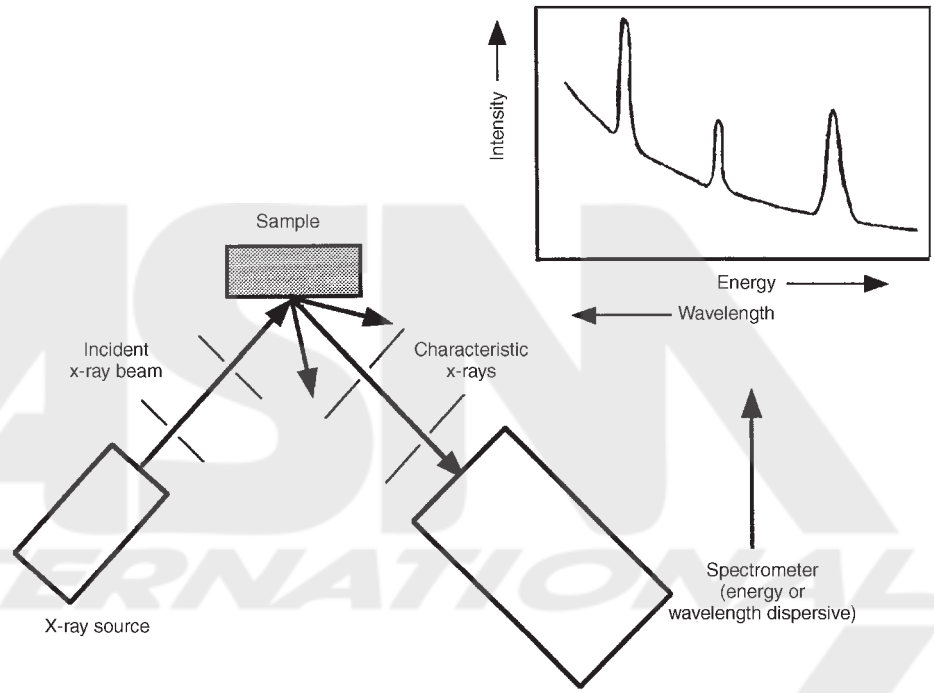


Fig. 2 Schematic of x-ray fluorescence spectrometer. X-rays emitted from the sample are analyzed to determine the characteristic energies or wavelengths of x-rays emitted and the intensities of the various characteristic energies. Source: Ref 1

of their inner shell electrons to higher energy levels. As the excited atoms relax, x-ray photons are emitted corresponding to the differences in the characteristic energy levels of the elements in the sample.

Qualitative analysis (determination of which elements are present) is done by comparing the energies of the x-rays emitted from the sample with the known characteristic x-ray spectra of each element (Fig. 3). Quantitative determination of the concentration of each element present is computed based on the intensities of the various characteristic x-ray energies, also shown in Fig. 3. Quantitative analyses can be most accurate by comparing the x-ray intensities from the unknown sample with their counterparts from a series of standard, similar, and known compositions. All modern instruments are equipped with computers to facilitate this calibration and measurement process. The use of progressively more powerful computer hardware and software has substantially decreased the need for standards with compositions tailored to specific classes of alloys. Many current analyses are done based only on pure element standards, using the computer to make composition-dependent corrections by iterative means.

Due to the dual particle and wave nature of electromagnetic radiation, a simple beam of x-rays can be thought of both as a wave with a characteristic wavelength and a stream of photons each having the same character-

istic energy. The photon energy is inversely related to the wavelength by the equation: photon energy = hc/λ, where h = Plank's constant, c = velocity of light, and λ= wavelength.

It is important to distinguish between x-ray energy and x-ray intensity. Energy is defined by the energies of the photons or the wavelength of the beam. Intensity is defined by the number of photons or the amplitude (height) of the wave. Of course, many x-ray beams are made up of numerous energies or wavelengths, not just one as previously described.

Wavelength Dispersive versus Energy Dispersive Detectors. The x-rays emitted from the sample in an XRF spectrometer are detected and analyzed in one of two ways: wavelength dispersive or energy dispersive analysis. In wavelength dispersive instruments, shown in Fig. 4, the emitted x-ray beam is directed onto one of several crystals that separates it into its component wavelengths by diffraction (similar to separating light into its component wavelengths by passing through a prism). An electronic counter is scanned over the angular range of the spectrometer, and a plot constructed of x-ray intensity versus wavelength (wavelength is calculated from the angle and characteristics of the diffracting crystal). Alternatively, the counter can be set to a series of predetermined wavelengths corresponding to the elements in the sample (presuming this is known or has already been determined), and the numbers of x-rays at each of these wavelengths counted for a specific length of time.

In energy dispersive instruments, the emitted x-ray beam is analyzed electronically, photon by photon, as illustrated in Fig. 5. The x-ray beam is directed into a semiconductor device (a lithium-drifted silicon crystal). As each x-ray photon enters the detector crystal, it creates numerous electron-

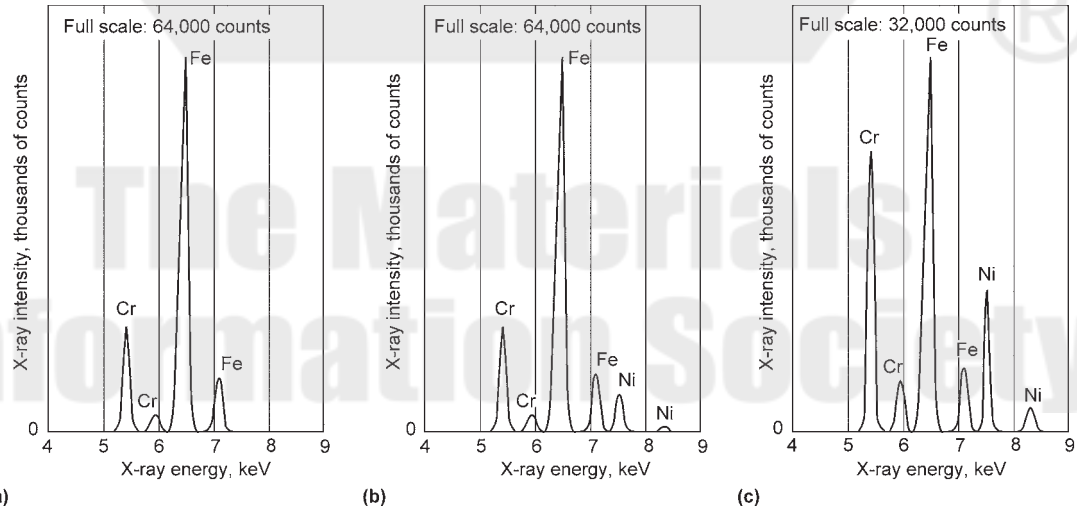


Fig. 3 X-ray fluorescence spectra of (a) Fe-16.4%Cr, (b) Fe-12.3%Cr-12.5%Ni, and (c) Fe-25.7%Cr-20.7%Ni. The iron, chromium, and nickel peaks occur at the same characteristic energies, but the intensities of the peaks increase with concentration. Courtesy of Sandia National Laboratories. Source: Ref 1

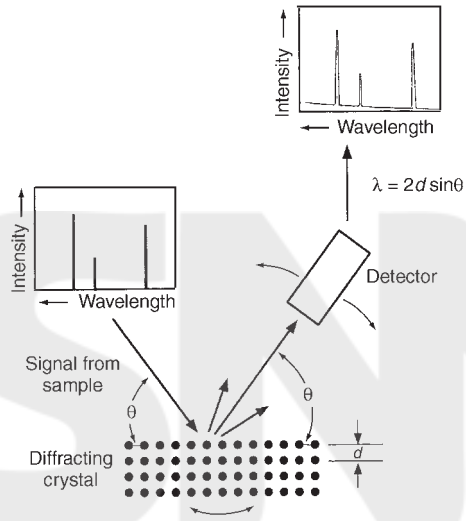


Fig. 4 Schematic of wavelength dispersive x-ray detector. Detector can mechanically scan a range of angles to produce a plot of intensity vs. wavelength, or it can be set at specific angles corresponding to the characteristic wavelengths of elements known to be in the sample, counting the x-ray intensity at each angle. Source: Ref 1

hole pairs as it expends its energy interacting with the atoms in the detector. These electrons and holes are collected and counted at the positive and negative bias sides of the detector. The energy of the photon is determined from the number of electron-hole pairs it creates (proportional to the energy of the photon). The detector electronics sense when each photon enters the detector and require that a photon's energy be analyzed before accepting input from any additional photons. Typically, several thousand photons are analyzed per second. As these energies are measured, a histogram of the numbers of photons counted corresponding with these energies is plotted on a cathode ray tube. The result is a digital plot of intensity versus energy, similar to what is obtained from wavelength dispersive spectrometers, as shown in Fig. 6.

Wavelength dispersive spectrometers (WDS) predate energy dispersive spectrometers (EDS), but each has inherent advantages. The primary advantage of energy dispersive systems is speed. They can collect a complete spectrum with several hundred thousand counts in approximately one minute. However, wavelength dispersive systems have superior energy resolution which can be important for separating signals from elements whose characteristic emission energies are very close to one another, as well as improved signal-to-noise ratios. As a result, energy dispersive systems are ideally suited for performing routine qualitative analyses, as well as quantitative analyses where speed is more important than the highest possible precision. On the other hand, difficult qualitative

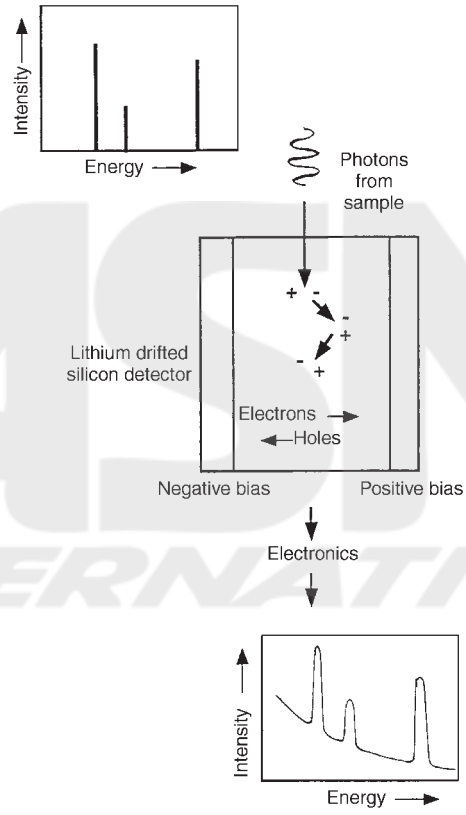


Fig. 5 Schematic of energy dispersive x-ray detector. Detector measures the energy of each incoming x-ray photon by counting the number of electron hole pairs it produces. A histogram is then developed and plotted of the x-ray energies of the many (typically tens to hundreds of thousands) photons measured during the counting period. Source: Ref 1

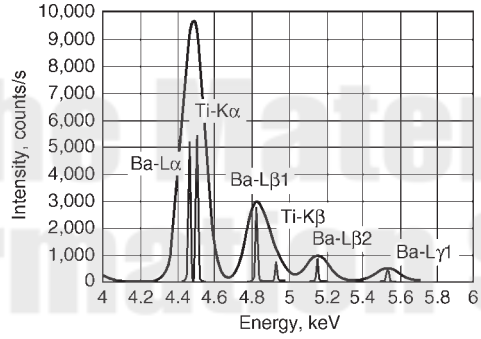


Fig. 6 Superimposed spectra of $BaTiO_3$ obtained from EDS and WDS systems (WDS spectrum replotted on energy scale, rather than wavelength). Note that the WDS spectrum has much sharper lines, thus enabling resolution of nearby peaks that overlap one another on EDS spectrum. Note also that the WDS spectrum has less background noise. Source: Ref 2

analyses and the most precise quantitative analyses can best be performed on the slower wavelength dispersive systems.

Fine Beam Instruments. Some XRF instruments are specifically designed to characterize compositional uniformity and to map compositional variations within a sample. Such instruments are typically capable of collimating the incident x-ray beam to smaller diameters as low as ~0.1 mm (~0.004 in.), thus enabling operator defined adjustment of lateral spatial resolution. In addition, the sample is usually mounted on an *x-y* stage that automatically translates it to the successive measurement points. The results of sequential analyses performed over a predefined trace on the sample are presented as a plot of composition versus position. Similarly, the results of many analyses performed over a predefined area are typically presented as a computer generated color coded two-dimensional map of composition versus position. This map provides a method for characterizing chemical inhomogeneities on a spatial resolution scale midway between the ~1 cm (~0.04 in.) range of bulk XRF and the ~1 μm scale of electron probe microanalysis, for example, characterizing segregation patterns in cross-sectioned ingots.

Capabilities of Related Techniques

Optical Emission Spectroscopy (OES) operates on the same atomic principles but bases its analyses on visible light, rather than x-rays. It has somewhat better sensitivity than XRF and better detection for some light elements, such as carbon. Some combinations of elements that exhibit interferences in the x-ray regime are free of interferences in the visible light regime and can be better analyzed by OES.

Combustion and vacuum fusion analysis is well suited to measuring gaseous impurities in metals.

Electron probe microanalysis (EPMA) operates on the same atomic principle as XRF, but it uses a focused electron beam to excite and generate x-rays in very small portions of the sample. Thus, it can be used to perform quantitative analyses on features as small as several micrometers and to generate quantitative elemental maps with several micrometer spatial resolutions. Further discussion of EPMA is covered in Chapter 8, "Metallography," in this book.

Optical Emission Spectroscopy (OES)

OES is capable of the detection and quantification of most elements except halogens, hydrogen, nitrogen, oxygen, and noble gases. The major use is the qualitative and quantitative elemental analyses of major, minor, and trace elements in metals and alloys.

The detection threshold of OES is on the order of tens of parts per million (ppm): 0.001 to 0.01%. The precision of quantitative analyses with

photographic instruments is ~5% relative or 0.05% absolute, whichever is greater, and for direct reading and charge coupled device (CCD) instruments, ~1% relative or 0.01% absolute, whichever is greater.

The limitations of OES are:

- Elements such as hydrogen, oxygen, nitrogen, halogens, and noble gases cannot be analyzed quantitatively
- Carbon and sulfur can only be measured in instruments equipped with vacuum chambers and in cases where the sample has not been powdered and mixed with these elements
- Some combinations of elements are difficult to determine because of overlapping energies in the visible light region

Both powders and bulk powders can be sampled. For powders, a ~1g (~0.002 lbs) sample is required; and, for bulk solids, about a 5 mm (0.2 in.) diameter surface spot and a sampling depth of ~100 μm are needed. Metal samples are typically ground to produce a flat uncontaminated surface for analysis. Most direct reading instruments can accommodate solid metal samples 10 cm (4 in.) or more in size. The technique is nondestructive except for the ~5 mm (~0.2 in.) diameter surface blemishes produced by the arc. Nonconductive samples are typically powdered and then mixed with a conductive low atomic number material, usually graphite. Samples as small as ~10 μg of powder (conductive or nonconductive) can be analyzed in this way.

Operating Principles

Physical Basis. Optical emission spectroscopy operates on the same atomic principles as XRF spectroscopy, except that analyses are based on visible light, rather than x-rays. Visible light is produced by transitions between electrons in the outer shells (far from the nucleus), while x-rays are produced by inner shell transitions. Photons of visible light have much lower energies than x-rays (1.5 to 4 electron volts for visible light, compared to 200 to 20,000 electron volts for x-rays), and correspondingly much longer wavelengths (~800 to ~300 nm for visible light, compared to 6 to 0.06 nm for x-rays).

The energies of the outer shell electrons on which OES is based can be substantially influenced by the surrounding atoms to which they are bonded to in solid samples. As a result, these bonds need to be completely broken for OES spectra to appropriately reflect the energies of the elements present in the sample. This break occurs by supplying sufficient energy to vaporize and decompose a portion of the sample into its component atoms, as well as to excite the atoms in this plasma.

Instrumentation. The external energy is frequently supplied by striking an electrical arc to the surface of the solid sample, as illustrated in Fig. 7. The arc vaporizes a small portion of the sample and ionizes the atoms,

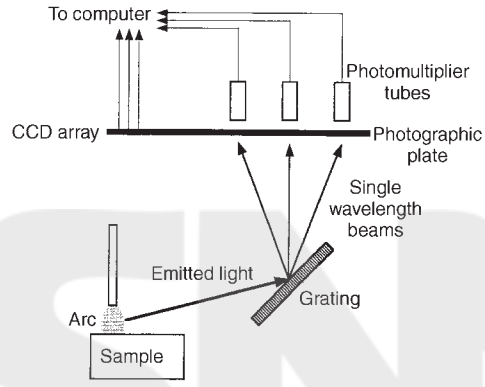


Fig. 7 Schematic of optical emission spectrometer. Light emitted from the vaporized and excited portion of the sample are analyzed to determine the characteristic wavelengths of light emitted and the intensities of the various characteristic wavelengths. Source: Ref 1

producing a plasma. Photons are emitted corresponding to the differences in the characteristic energy levels of the elements in the plasma. The visible light portion is analyzed by passing it through a grating to separate it into its various component wavelengths (similar to separating white light into its component wavelengths by passing it through a glass prism). The resulting spectrum is recorded and compared to the spectra of known elements to determine what elements are present. The concentrations of each element present are deduced from the intensities (number of photons) corresponding to each characteristic wavelength. Quantitative determination of the concentration of each element present is done by comparing the intensities from the unknown sample with their counterparts from a series of standards of known concentration.

Photographic Instruments. Two methods have historically been used for recording the spectra and analyzing the data, but both of these are currently being replaced by a third method based on newer technology. In photographic emission spectroscopy, the various emitted wavelengths and intensities coming from the grating are directly recorded on a photographic plate. The elements present and their concentrations are deduced by direct visual comparison of the spectrum from the sample and spectra from standards of known composition. This process can be made more quantitative by using a densitometer to read the plates and compare the spectra. Photographic optical emission spectroscopy is readily used for both qualitative and quantitative analysis. Major and minor alloying additions can be readily detected and quantified. Trace elements can typically be detected and quantified down to the range of 10 to 100 ppm.

Direct Reading Instruments. Direct reading emission spectrometers have more often been used in cases where particular combinations of elements must be routinely and frequently quantified, such as in measuring

the compositions of heats being produced by an aluminum ingot production facility. In these instruments, photomultiplier tubes are set up at the wavelength positions corresponding to each element of interest. These tubes measure the intensities of light obtained from the sample at each predetermined characteristic wavelength and input the results to a computer. The computer then compares these intensities to the corresponding intensities from various standards of known composition (which are already stored in its memory) and calculates the composition of the sample in a matter of seconds. This facilitates rapid quantitative analysis of samples of unknown but similar composition. Because they are only set up to analyze for specific preselected elements and over commonly encountered composition ranges, direct reading emission spectrometers are not useful for qualitative or quantitative analysis of broader ranges of samples.

Instruments with Charge Coupled Device (CCD) Detectors. Both of the previously described methods of detection are now being supplanted by CCD detectors (a high-resolution array of solid state light detectors). These detectors take the place of photographic plates and electronically record both the location or wavelength and intensity or brightness of the light emitted from the sample. The output of the CCD detector is inputted to a computer, which constructs a histogram of intensity versus wavelength. The computer also facilitates comparison of the observed wavelengths with the known characteristic wavelengths of each element, thus providing for qualitative determination of which elements are present in the sample. In addition, the computer can make quantitative determinations of the concentrations of each element present from the intensities of the emitted light at various characteristic energies. These determinations can be made based on comparison with calibrated standards (measured by the instrument and stored in the computer memory) or based on pure element standards. Analyses based on calibrated standards are typically used in cases where greater accuracy and precision are required. In essence, then, an instrument equipped with a CCD detector is like a direct reading instrument with a semi-infinite number of detectors positioned at every wavelength of possible interest, which provides a rapid and powerful capability for both qualitative and quantitative analysis.

Capabilities of Related Techniques

Inductively coupled plasma atomic emission spectroscopy (ICP/AES) operates on the same atomic principle. However, solid samples are dissolved into liquid solutions that are then aspirated into an argon plasma. This process provides greater flexibility, as liquid standards can be made of essentially any combination of elements, including combinations that cannot be obtained in solid form. Inductively coupled plasma can be used for both qualitative and quantitative elemental analysis. However, its detection limits are lower than those of other OES methods. Because the

solid sample must be dissolved and diluted in a liquid, the effective detection limits for analyzing solid samples are similar.

Inductively Coupled Plasma Mass Spectroscopy (ICP/MS). This combined technique provides better detection limits for trace elements, in some cases down to the range of parts per trillion. The dissolved sample is aspirated into the plasma and ionized. Ions from the plasma are then input to a mass spectrometer that determines which elements are present in the plasma. The increased sensitivity of the mass spectrometer provides for lower detection limits, typically in the range of parts per billion.

Atomic absorption spectroscopy (AAS) operates on the same atomic principle as OES, but it measures the intensity of light absorbed by the liquid sample aspirated into a flame or graphite furnace, rather than the light emitted. Flame AAS has similar sensitivity to OES. Graphite furnace AAS exhibits better sensitivity for trace elements, similar to ICP/MS, but is simpler and less expensive to set up and operate. Because only one element can be measured at a time, single wavelength light sources are used for each element, which makes AAS inappropriate for qualitative analysis.

X-ray fluorescence spectroscopy (XRF) operates on similar physical principles to OES, but it excites the sample using x-rays rather than thermal energy and analyzes the x-rays emitted from the sample rather than the visible light. X-ray fluorescence spectroscopy is completely nondestructive and can be used for both qualitative and quantitative analysis. It provides a good complement to OES, because the interferences or overlapping energies in the x-ray regime are different from those in the visible light regime. It is not as sensitive as OES for analyzing trace elements. Detection limits are typically in the range of 100 to 1000 ppm.

Combustion and Inert Gas Fusion Analysis

Combustion and inert gas fusion analysis is used to conduct quantitative analysis of the amounts of carbon, sulfur, and dissolved gases in metals.

The threshold sensitivity is in the vicinity of 50 ppm (0.005%) for sulfur, 10 ppm (0.001%) for carbon, 1 ppm for oxygen, nitrogen, and hydrogen. The precision for quantitative analysis in sulfur is, ~20% relative or 0.005% absolute, whichever is greater; carbon, ~5% relative or 0.001% absolute, whichever is greater; oxygen, nitrogen, and hydrogen, ~5% relative or 0.0001% absolute, whichever is greater.

The amount of material sampled is typically ~1 g (0.002 lbs). Solid samples in the range of 0.1 g to several grams are usually machined to fit into the sample cups of the instrument. It is desirable to use regular shapes with small ratios of surface area to volume in order to minimize the contributions of surface adsorbed gases to the results. Samples should be cleaned immediately prior to analysis.

The limitations are:

- Samples containing highly stable nitrides or oxides require special treatment.
- Metals with low boiling points require special treatment.

Operating Principles

A schematic diagram of the similar high temperature combustion and inert gas fusion processes is shown in Fig. 8. Small samples of known weight are heated to very high temperatures. The elements of interest are driven off as either elemental gas or gaseous oxidation products. These gaseous products are then separated and detected, permitting quantification of their concentrations in the original samples.

In the case of combustion analysis for carbon or sulfur, the sample is induction or resistance heated to ~1400 °C (2550 °F) in oxygen, which causes the sample to be completely oxidized. The metal oxides are left as solid, but the carbon and sulfur form CO, CO_2, and SO_2, which are liberated as gases. These gases are then passed through a series of traps, absorbers, and converters to separate them and remove interfering elements. They are then quantified using detectors based on either thermal conductivity or absorption of infrared light.

In the case of inert gas fusion for oxygen, nitrogen, or hydrogen, the sample is heated by either induction or by passing a high current through it to ~3000 °C (5450 °F) in an inert gas. The dissolved gases are driven off at this extremely high temperature. In some cases, these gases combine

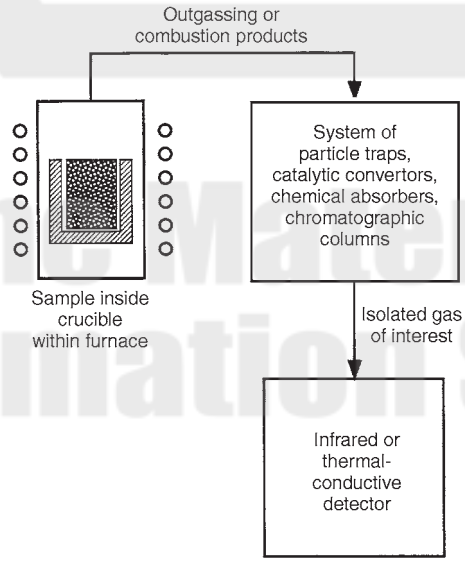


Fig. 8 Schematic of combustion/inert gas fusion apparatus. Source: Ref 1

with other elements in the system; for example, oxygen reacts with carbon to form CO or CO_2. The liberated gases and/or reaction products are then passed through a series of particle traps, catalytic converters, gas chemical absorbers, and chromatographic columns to separate and purify them. They are then quantified using detectors based on either thermal conductivity or absorption of infrared light, as in combustion analysis.

Capabilities of Related Techniques

X-ray fluorescence spectroscopy (XRF) can analyze nondestructively for carbon and sulfur but with higher detection limits and reduced precision.

Optical emission spectroscopy (OES) can analyze less destructively for carbon and sulfur, but with higher detection limits and reduced precision.

Hot extraction high vacuum analysis is similar to inert gas fusion, but gases are liberated at lower temperatures (without destroying sample). Gas evolution can be monitored as a function of time or temperature, thus permitting separation of internally dissolved and surface adsorbed gases.

Surface Analysis

A number of methods can be used to obtain information about the chemistry of the several atomic layers of samples of metals, as well as of other materials, such as semiconductors and various types of thin films. Of these methods, the scanning Auger microprobe is the most widely used. This instrument and method will be described in some detail, and the operation and capabilities of several other surface analysis methods will be more briefly described in comparison to it.

Scanning Auger Microprobe (SAM)

The scanning Auger microprobe is basically a scanning electron microscope (SEM) with two additional features:

- An Auger electron detector replaces an x-ray detector. The Auger detector is used to measure the energies of Auger electrons emitted from the sample. These characteristic energies enable identification of the elements present in the first few atomic layers of the surface. The concentrations of each element present can also be determined by the number of electrons detected at each of the characteristic energies. All elements except hydrogen and helium can be identified and analyzed in this way.
- An in situ ion milling capability provides for gradual removal of surface layers, thereby permitting depth profiling of elemental compositions within about 1 μm of the surface.

These capabilities make the SAM well suited for the following types of applications:

- Identification and mapping of light elements (atomic numbers 3 to ~9) that are difficult to detect using SEM or electron probe microanalysis (EPMA)
- Elemental characterization of surface contaminants
- Depth profiling of elemental compositions within ~1 μm of the surface (this is particularly widely used in microelectronics applications)

The spatial resolution of secondary electron imaging of surface topography is ~10 nm (same as SEM), and the resolution of Auger electron characterization of elemental chemistry is 10 to 20 nm at a sampling depth of ~1 nm. The threshold sensitivity is ~0.5%, and the precision for quantitative analysis is ~10% relative or 0.5% absolute, whichever is greater.

Samples up to ~2.5 cm (~0.10 in.) diameter and 0.5 cm (0.20 in.) thick can be accommodated by most SAMs; larger samples can be accommodated in instruments designed for this purpose. Provisions must be made for charge to bleed off. Ideal samples are electrically conductive and must be free of fingerprints, oils, and other high vapor pressure materials. Flat samples are preferred, but rough samples can also be accommodated.

Limitations of SAM are:

- Cannot detect hydrogen or helium
- Quantitative analyses are typically lower in quality than those of EPMA

Operating Principles

Instrumentation. As noted above, the scanning Auger microprobe is essentially a SEM to which an Auger electron detector and an ion miller have been added as shown in Fig. 9 An electron beam is produced and focused to a small spot on the sample surface. This spot can be rastered across an operator defined area of the surface or stopped and moved to a particular location of interest. The beam penetrates the sample and interacts with the atoms in the first ~1 μm, exciting atoms and producing secondary electrons, exactly like a SEM. A secondary electron detector provides the capability to image the surface and locate areas of particular interest, as in a scanning electron microscope. However, the primary tool for chemical analysis is the Auger electron detector.

Physical Basis. Auger electrons are emitted as the excited atoms relax. In a sense, they are the complements of the characteristic x-rays that are used for chemical characterization in x-ray fluorescence (XRF), SEM, EPMA, and transmission electron microscopy (TEM). When atoms become excited by electrons being ejected from their inner shells, electrons from higher energy shells fill these vacated sites. This process always results in the release of energy equal to the energy difference between the

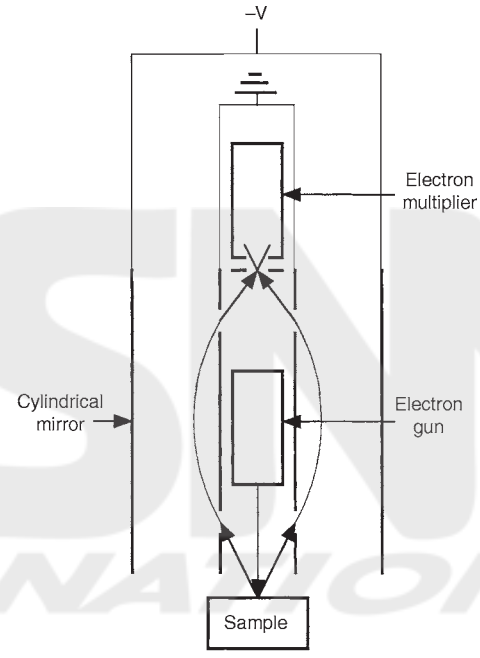


Fig. 9 Schematic of a scanning Auger microprobe. Source: Ref 3

donor and acceptor levels. However, emission of characteristic x-rays is only one of the mechanisms by which this energy can be released. Another common mechanism is by the release of an Auger electron.

An Auger electron is an electron from one of the outer shells that is ejected from the atom with kinetic energy equal to the energy released by the relaxation event minus the energy required to remove the Auger electron from its orbit, as illustrated in Fig. 10. Because both energies associated with the relaxation events and the binding energies of the outer shell electrons provide characteristic *fingerprints* for each element, so do their differences, the energies of Auger electrons. Hence, the energies of Auger electrons can also be detected and used to identify which elements are in the portion of the sample being excited by the incident electron beam. These characteristic Auger electrons typically have energies of tens to thousands of electron volts.

The tendency for excited atoms to relax by Auger electron production versus x-ray photon emission increases with decreasing atomic number. Elements with atomic numbers less than ~7 produce few characteristic x-rays but many Auger electrons (except for hydrogen and helium). As a result, SAM is commonly used for microstructural detection and quantification of such elements. However, higher atomic number elements produce more x-rays, so these elements are typically detected and quantified using SEM or EPMA. Although when analysis of the first few atomic lay-

ers is desired, SAM provides elemental analyses corresponding to this very near surface region.

While Auger electrons are generated throughout the beam-sample interaction volume, most of these dissipate some or all of their characteristic energies by interacting with the electrons belonging to other atoms in the sample. The only Auger electrons that escape the sample with their original characteristic energies are those generated within a few atom layers of the sample's surface. If the energies of all emitted electrons are detected and analyzed, then a graph similar to Fig. 11 is obtained. The lowest energy range is dominated by secondary electrons, the highest energy range is dominated by backscattered electrons, and the mid-energy range is dominated by Auger electrons, nearly all of which have had their charac-

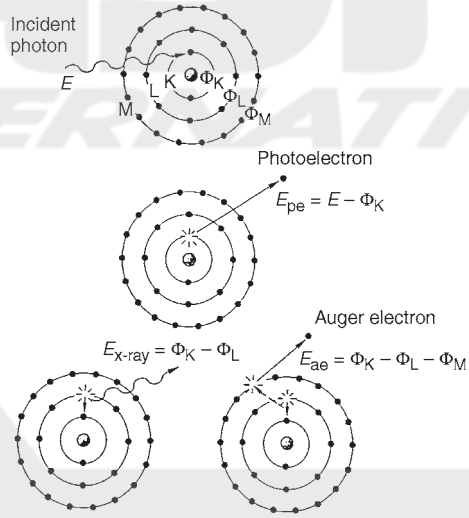


Fig. 10 Comparison of production of x-rays and Auger electrons. Source: Ref 4

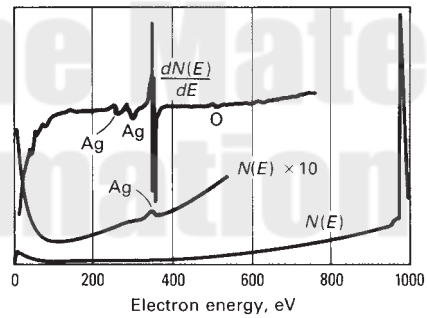


Fig. 11 Electron energy distribution from silver sample. Differentiated signal most clearly reveals the peak corresponding to Auger electrons that were produced very near the surface and exited the sample prior to interacting with other atoms. Source: Ref 5

teristic energies reduced by interactions with the sample. But if Fig. 11 is examined closely, small signals can be found at particular energies. These represent the characteristic energies of the *undisturbed* Auger electrons that were generated by atoms at the surface or within a few atomic distances below the surface.

Electron Collection and Energy Measurement. The energies of the emitted electrons are usually measured using a cylindrical *mirror* that has a variable negative potential applied to it, as shown in Fig. 9. As electrons enter the inlet aperture and pass through the analyzer chamber, the negative bias on the wall of the chamber repels them and causes them to travel in curved paths. The curvature of this path varies inversely with the kinetic energy of each electron. The paths of electrons with low kinetic energy are more easily deflected than those paths of electrons with high kinetic energy. This information provides a means of measuring the energy distribution of the electrons emitted from the sample. An electron detector is mounted near the exit aperture of the cylindrical mirror, and the negative bias applied to the mirror is gradually increased. The numbers of electrons entering the detector is counted as a function of mirror bias. This information enables the energy distribution of the electrons to be plotted.

The portion of the signal corresponding to the *undisturbed* Auger electrons is very small compared with the signal resulting from backscattered and Auger electrons whose energies have been reduced by interactions within the sample. This situation is typically overcome by differentiating the signal and plotting dN/dE versus E, as shown in Fig. 11. Because the Auger electrons typically originate in the outer electron shells, their energies are somewhat affected by bonding between atoms. These small energy shifts, which can frequently be discriminated by the energy analyzer, provide the ability to determine some information about the elements to which the atoms of interest are bonded.

Scanning Auger microprobe results are often presented as secondary electron images with accompanying Auger electron spectra identifying the elements present in particular features of interest. Low atomic number elements that cannot be detected by SEM and EPMA are readily detected in Auger spectra. An example of the use of Auger electrons to detect low atomic number elements with high spatial resolution is shown in Fig. 12. Alternatively, the detector can be set to the energy associated with a particular element or compound of interest and the electron beam rastered over the surface, resulting in a map indicating the areas of high concentration of this material (Fig. 13).

An ion sputtering gun is also incorporated into the chamber of the instrument and can be used to progressively remove thin layers of material from the surface of the sample. This removal provides the opportunity to perform Auger elemental measurements at various distances from the original sample surface. The practical limit of such depth profiling is ~1 μm. An example of this depth profiling capability is shown in Fig. 14.

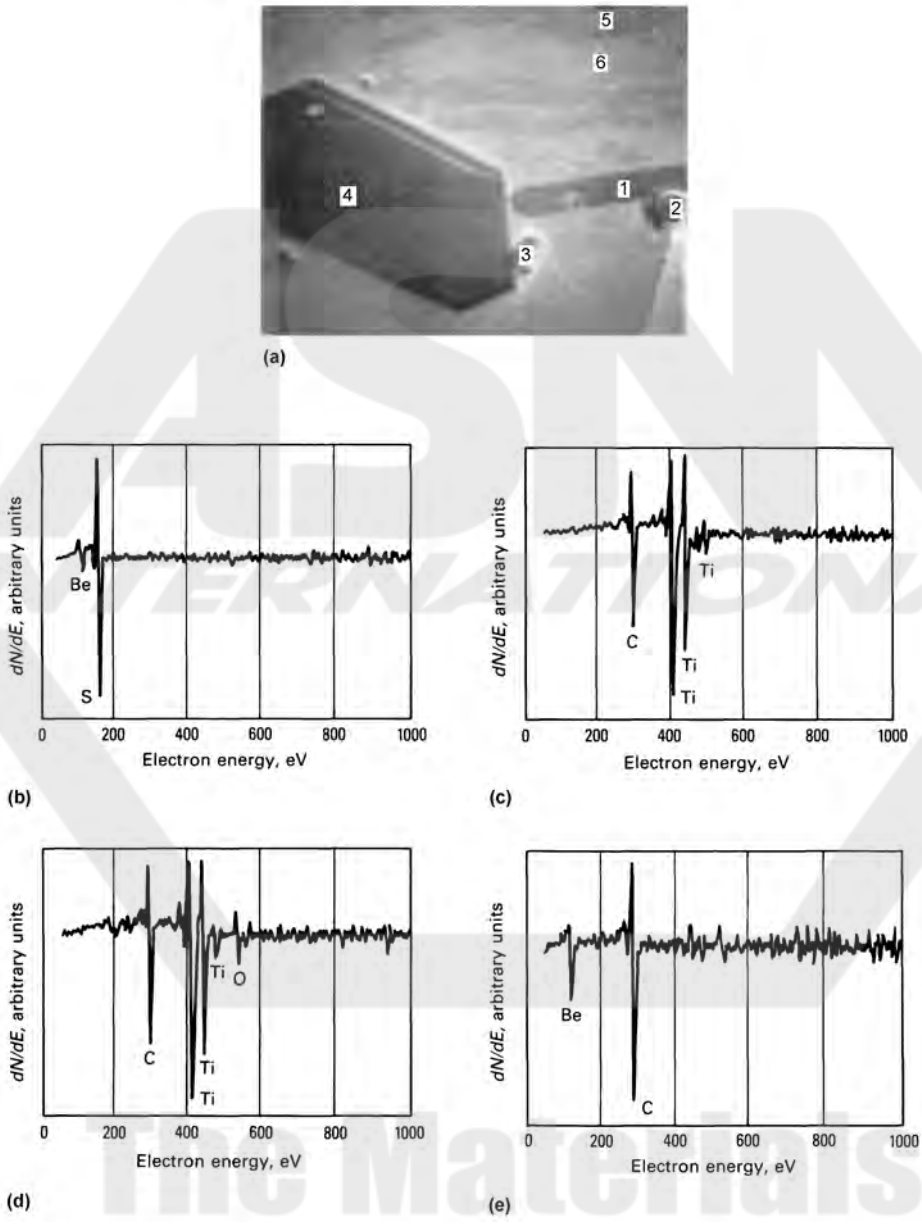


Fig. 12 Scanning Auger identification of elements, including some of low atomic number, present in several phases in a copper-beryllium alloy. (a) Secondary electron image showing inclusions. (b-e) Auger spectra obtained from the indicated microstructural features. (b) The long rod shaped precipitate (point 1) is a beryllium sulfide. (c) The small round precipitate (point 2) is a titanium carbide. (d) The small irregular precipitate (point 3) is also a titanium carbide. (e) The large blocky angular precipitate (point 4) is a beryllium carbide. Source: Ref 5

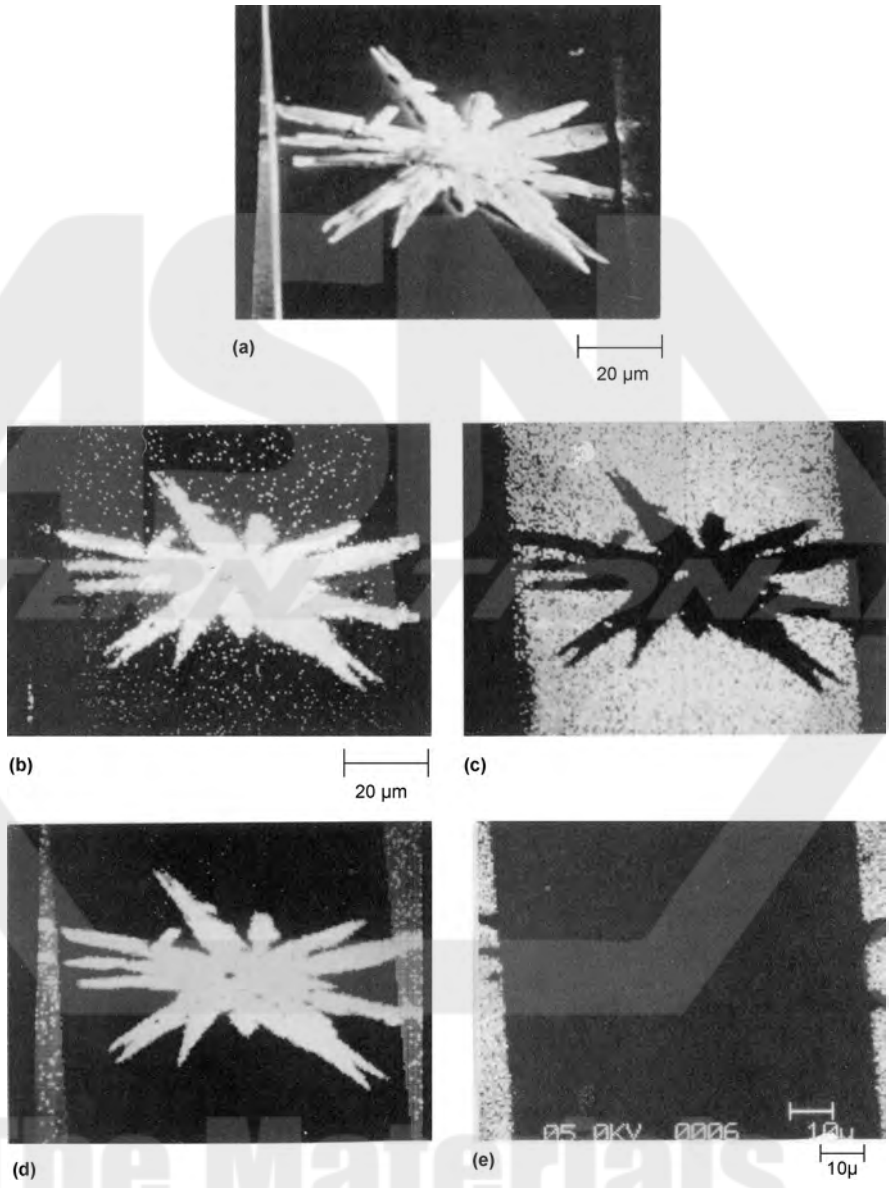


Fig. 13 Scanning Auger mapping of elements, including some of low atomic number, in a foreign particle on an integrated circuit. Note also the ability to distinguish between elemental silicon and silicon oxide due to bonding effects on Auger energies. (a) Secondary electron image of particle. (b–e) Auger maps showing locations of silicon oxide, elemental silicon, oxygen, and aluminum, respectively. Source: Ref 5

Related Surface Analysis Techniques

Nonsurface Specific Methods

Nonsurface specific methods include scanning electron microscopy (SEM), electron probe microanalysis (EPMA), and transmission electron microscopy (TEM).

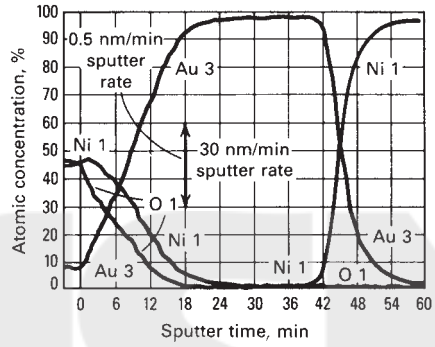


Fig. 14 Scanning Auger depth composition profile obtained from a nickel-rich area of a gold-nickel-copper metallization surface. Source: Ref 5

Scanning Electron Microscopy, Electron Probe Microanalysis (SEM, EPMA). These methods are better for combined imaging and elemental analysis of elements with higher atomic numbers. However, SEM and EPMA cannot readily analyze for low atomic number elements (less than ~7 to 11), nor do they have either depth profiling capabilities or surface specific analytical capabilities. Further discussion of SEM and EPMA, is covered in Chapter 8, "Metallography," in this book.

Transmission electron microscopy (TEM) has better spatial resolution for imaging and chemical analysis, but an EDS system cannot readily analyze for low atomic number elements (below ~7). The presence of lower atomic number elements can sometimes be inferred by electron diffraction if these elements are present in compounds that can be identified based on their interplanar spacings. Further discussion of TEM is covered in Chapter 8, "Metallography," in this book.

Surface-Specific Methods

A number of other techniques are frequently used to characterize the chemistries of the top one to five atomic layers of materials. The following provides brief summaries of two of the methods that are frequently used in metallurgical studies and comparisons of their capabilities with those of the scanning Auger microprobe.

Secondary ion mass spectroscopy (SIMS) directs a finely focused beam of energetic ions onto the sample surface, then it collects and analyzes the ionized atoms or clusters of atoms ejected from the sample surface by this beam. Information can be obtained with lateral spatial resolution of 100 to 500 nm. The ions removed from the surface are identified by a highly sensitive mass spectrometer. This identification provides for very sensitive detection of many elements, often in the range of parts per billion. In addition, it enables analysis for very low atomic number elements; including hydrogen (SIMS is the only method able to detect hydrogen with microscopic spatial resolution). The primary ion beam can be rastered over the surface, providing for high sensitivity elemental mapping.

Because it removes material from the surface, it also provides for depth profiling.

X-ray photoelectron spectroscopy (XPS) directs a single energy x-ray beam onto the surface. This beam penetrates 10 to 100 μm into the sample, interacting with atoms and ejecting photoelectrons from their inner shells. The energies of these photoelectrons are equal to the energy of the x-ray photons minus the characteristic electron binding energies. Many of the photoelectrons lose some or all of their energy in interactions with other atoms, but a few that are generated very close to the surface exit the sample undisturbed. The photoelectrons are collected and their energies analyzed using a device similar to the cylindrical mirror in a scanning Auger microprobe (SAM). Analysis of the energies of the photoelectrons permits identification of the elements in the top few atomic layers. The excellent energy resolution of the analyzer enables it to discriminate the very small shifts in energy that result from bonding of the atoms of interest to other surrounding atoms. Hence, XPS is capable of providing information on surrounding atoms and chemical bonding. X-ray photoelectron spectroscopy does not utilize a fine incident beam; therefore, it does not provide images or chemical information with high lateral spatial resolution. It is generally not as sensitive as SIMS, but it is very useful for detecting some elements for which SIMS is not very sensitive. An ion sputtering capability is generally available to facilitate depth profiling. In general, XPS is most extensively used to obtain surface analyses with chemical bonding sensitivity.

ACKNOWLEDGMENT

This chapter was adapted from Bulk Elemental Analysis and Surface Analysis, both by K.H. Eckelmeyer in *Metals Handbook Desk Edition*, Second Edition, ASM International, 1998.

REFERENCES

1. K.H. Eckelmeyer, Bulk Elemental Analysis, *Metals Handbook Desk Edition*, 2nd ed., ASM International, 1998, p 1410–1411
2. C. Brundle, C. Evans, and S. Wilson, *Encyclopedia of Materials Characterization*, Butterworth-Heinemann, 1992, p 128
3. K.H. Eckelmeyer, Surface Analysis, *Metals Handbook Desk Edition*, 2nd ed., ASM International, 1998, p 1433–1436
4. R. Jenkins, R. Gould, and D. Gedcke, *Quantitative X-Ray Spectrometry*, Dekker, 1981, p 16
5. *Materials Characterization*, Vol 10, *ASM Handbook*, ASM International, 1986

CHAPTER 8

Metallography

THE METHODS AND EQUIPMENT described in this chapter cover the preparation of specimens for examination by light optical microscopy (LOM). It is assumed that the specimen or specimens being prepared are representative of the material to be examined. However, random sampling, as advocated by statisticians, can rarely be performed by metallographers. Instead, metallographers are usually restricted to inspecting areas of interest or systematically chosen test locations based on sampling convenience. In failure analysis, specimens are usually removed to study the origin of the failure, to examine highly stressed areas, and to examine secondary cracks.

All sectioning processes produce damage; some methods, such as flame cutting, produce extreme amounts of damage. Traditional laboratory sectioning procedures using abrasive cut-off saws introduce minor damage that varies with the material being cut and its thermal and mechanical history. This damage must be removed if the true structure is to be examined. However, because abrasive grinding and polishing steps also produce damage, where the depth of damage decreases with decreasing abrasive size, the preparation sequence must be carefully planned and performed. Otherwise, preparation induced artifacts will be interpreted as structural elements. The characteristics of a properly prepared specimen are:

- Deformation induced by sectioning, grinding, and polishing is removed or shallow enough to be removed by the etchant
- Coarse grinding scratches are removed; fine polishing scratches are tolerated in routine metallographic studies
- Pullout, pitting, cracking of hard particles, and smear are avoided
- Relief (i.e., excessive surface height variations between structural features of different hardness) is minimized
- The surface is flat particularly at edges, if they are to be examined, and at coated surfaces to permit examination at high magnifications

- Specimens are cleaned adequately between preparation steps, after preparation, and after etching

Preparation of metallographic specimens generally requires five major operations: sectioning, mounting (optional), grinding, polishing, and etching (optional).

Sectioning

Many metallographic studies require more than one specimen. For example, a study of deformation in wrought metals usually requires two sections–one perpendicular and the other parallel to the direction of deformation.

Sampling. Bulk samples for sectioning may be removed from larger pieces or parts using methods such as core drilling, band or hack sawing, flame cutting, etc. However, when these techniques are used, precautions must be taken to avoid alteration of the microstructure in the area of interest. Laboratory abrasive wheel cutting is recommended to establish the desired plan of polish.

Abrasive Wheel Cutting. By far, the most widely used sectioning devices in metallographic laboratories are abrasive cut-off machines. All abrasive wheel sectioning should be done wet. An ample flow of water, with a water soluble oil additive for corrosion protection, should be directed into the cut. Wet cutting will produce a smooth surface finish and, most importantly, will guard against excessive surface damage caused by overheating. Abrasive wheels should be selected according to the recommendations of the manufacturer. Specimens must be fixtured securely during cutting, and cutting pressure should be applied carefully to prevent wheel breakage.

Mounting of Specimens

The primary purpose of mounting metallographic specimens is for convenience in handling specimens of difficult shapes or sizes during the subsequent steps of metallographic preparation and examination. A secondary purpose is to protect and preserve extreme edges or surface defects during metallographic preparation. The method of mounting should in no way be injurious to the microstructure of the specimen. Mechanical deformation and heat are the most likely sources of injurious effects.

Clamp Mounting. Clamps have been used for mounting metallographic cross-sections in the form of thin sheets. Several specimens can be clamped conveniently in sandwich form. This method is quick and convenient for mounting sheet type specimens; and when done properly, edge retention is excellent. There is no problem with seepage of fluids from

crevices between specimens. The outer clamp edges must be beveled to minimize damage to polishing cloths. If clamps are improperly used so that gaps exist between specimens, fluids and abrasives can become entrapped and will seep out obscuring edges. The problem can be minimized by proper tightening of clamps, by use of plastic spacers between specimens, or by coating specimen surfaces with epoxy before tightening.

Compression Mounting. The most common mounting method uses pressure and heat to encapsulate the specimen with a thermosetting or thermoplastic mounting material. Common thermosetting resins include phenolic, such as bakelite and diallyl phthalate, while methyl methacrylate is the most common thermoplastic mounting resin. Both thermosetting and thermoplastic materials require heat and pressure during the molding cycle; but after curing, mounts made of thermoplastic resins must be cooled to ambient under pressure, while mounts made of thermosetting materials may be ejected from the mold at the maximum molding temperature. Thermosetting epoxy resins provide the best edge retention of these resins and are less affected by hot etchants than phenolic resins. Mounting presses vary from simple laboratory jacks with a heater and mold assembly to full automated devices.

Cold mounting materials require neither pressure nor external heat and are recommended for mounting specimens that are sensitive to heat and/or pressure. Acrylic resins are the most widely used castable resin due to their low cost and fast curing time; however, shrinkage is somewhat of a problem. Epoxy resins, although more expensive than acrylics, are commonly used because epoxy will physically adhere to specimens and can be drawn into cracks and pores, particularly if a vacuum impregnation chamber is employed. Hence, epoxies are very suitable for mounting fragile or friable specimens and corroded or oxidized specimens. Most epoxies are cured at room temperature, but curing times can be as long as 6 to 12 hours. Some epoxies can be cured at slightly elevated temperatures in less time. Hard filler particles have been added to epoxy mounts for edge retention, but this addition is really not a satisfactory solution.

Taper sectioning (mounting) generally is regarded as a special mounting technique. It enables the metallographer to examine in greater detail the immediate subsurface structure or surface topography of a specimen. Microhardness determinations and thickness measurements of thin surface coatings or diffusion zones can be performed on taper sectioned specimens. Taper sectioning (Fig. 1) is accomplished by establishing a plane of polish at a small angle to the surface of the specimen.

Edge preservation is a long standing metallographic problem and many "tricks" have been promoted; most pertaining to mounting, but some to grinding and polishing. These methods include the use of backup material in the mount, the application of coatings to the surfaces before mounting, and the addition of a filler material to the mount. Plating of a

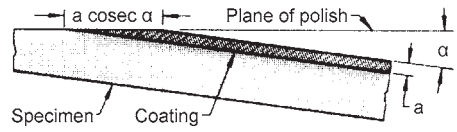


Fig. 1 Schematic of taper sectioning (mounting), as applied to a coated specimen. Taper magnification equals the cosecant of taper angle α. Source: Ref 1

compatible metal on the surface to be protected (electroless nickel has been widely used) is generally considered to be the most effective procedure.

However, introduction of new technology has greatly reduced edge preservation problems. First, use of mounting presses, which cool the specimen to near ambient temperature under pressure, produces much tighter mounts. Gaps that form between specimen and resin are a major contributor to edge rounding. Second, use of semiautomatic and automatic grinding/polishing equipment increases surface flatness and edge retention opposed to manual (hand) preparation. Third, the use of harder, woven or nonwoven, napless surfaces for polishing with diamond abrasives, rather than softer cloths, such as canvas, billiard, and felt, maintains flatness. Final polishing using low nap cloths for short times introduces very little rounding compared to use of higher nap, softer cloths.

Grinding

Grinding should commence with the finest grit size that will establish an initially flat surface and remove the effects of sectioning within a few minutes. An abrasive grit size of 180 or 240 grit is coarse enough to use on specimen surfaces sectioned by an abrasive cut-off wheel. Hack sawed, band sawed, or other rough surfaces usually require abrasive grit sizes from 120 to 180 grit. The abrasive used for each succeeding grinding operation should be one or two grit sizes smaller than that used in the preceding operation. A satisfactory fine grinding sequence might involve grit sizes of 240, 320, 400, and 600 grit. This sequence is known as the *traditional* approach.

As in abrasive wheel sectioning, all grinding should be done wet provided that water has no adverse effects on any constituents of the microstructure. Wet grinding: (a) minimizes loading of the grinding abrasive, and (b) prevents the sample from getting hot.

Each grinding step, while producing damage itself, must remove the damage from the previous step. The depth of damage decreases with the abrasive size but so does the metal removal rate. For a given abrasive size, the depth of damage introduced is greater for soft materials than for hard materials.

Besides SiC paper, a number of other options are available to circumvent their use. One option, used chiefly with semiautomatic and automatic systems, is to grind a number of specimens placed in a holder simultaneously using a conventional grinding stone generally made of coarse grit alumina. This step, often called *planar grinding*, has the second goal of making all of the specimen surfaces coplanar. This process requires a special purpose machine because the stone must rotate at a high speed, ≥1500 rpm, to cut effectively. The stone must be dressed regularly with a diamond tool to maintain flatness.

Other materials have also been used both for the planar grinding stage or, afterwards, to replace SiC paper. For very hard materials such as ceramics and sintered carbides, two or more metal bonded or resin bonded diamond disks with grit sizes from about 70 to 9 μm can be used. An alternate type of disk has diamond particles suspended in a resin applied in small blobs, or spots, to a disk surface. These disks are available with diamond sizes from 120 to 6 μm. Another type of disk available in several diamond sizes uses diamond attached to the edges of a perforated, screen like metal disk. Another approach uses a stainless steel woven mesh *cloth* on a platen charged with coarse diamond, usually in slurry form, for planar grinding. Once planar surfaces have been obtained, there are several single step procedures available for avoiding the finer SiC papers. These processes include the use of platens, woven polyester thick cloths, or rigid grinding disks. With each of these, a coarse diamond size, most commonly 9 μm, is used.

Grinding Media. The grinding abrasives commonly used in the preparation of metallographic specimens are silicon carbide (SiC), aluminum oxide (Al_2O_3), emery (Al_2O_3-Fe_3O_4), composite ceramics, and diamond. All except diamond are generally bonded to paper or cloth backing materials of various weights in the form of sheets, disks, and belts of various sizes. Limited use is made of grinding wheels consisting of abrasives embedded in a bonding material. The abrasives may be used also in powder form by charging the grinding surfaces with loose abrasive particles or with abrasive in a premixed slurry or suspension.

Grinding Equipment. Although it is rarely used in industry, students still use stationary grinding paper that is supplied in strips or rolls. The specimen is slid against the paper from top to bottom. Grinding in one direction is usually safer than grinding in both directions. While this can be done dry for certain delicate materials, water is usually added to keep the specimen surface cool and to carry the swarf away.

Belt grinders are usually present in most laboratories. They are mainly used to remove burrs from sectioning; to round edges that need not be preserved for examination; to flatten cut surfaces to be macroetched; or to remove sectioning damage. Generally, only very coarse abrasive papers with 60 to 240 grits are used. Most grinding work is done on rotating wheels; that is, a motor driven platen upon which the SiC paper is attached (Fig. 2).

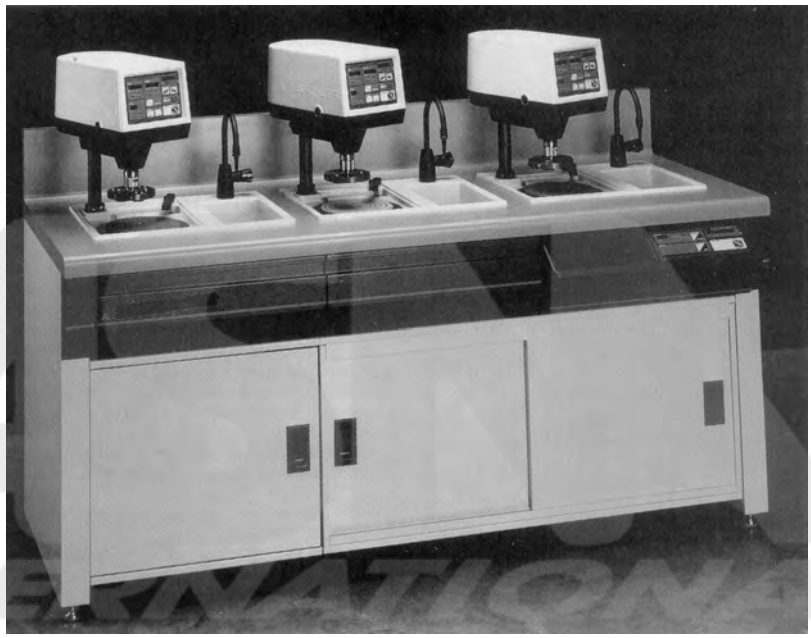


Fig. 2 Laboratory flush mounted semiautomatic grinder/polisher system. Source: Ref 1

Lapping is an abrasive technique in which the abrasive particles roll freely on the surface of a carrier disk. During the lapping process, the disk is charged with small amounts of a hard abrasive, such as diamond or silicon carbide. Lapping disks can be made of many different materials; cast iron and plastic are most commonly used. Lapping produces a flatter specimen surface than grinding, but it does not remove metal as in grinding. Consequently, lapping is not commonly employed in metallography. Some platens, referred to as laps, are charged with diamond slurries. Initially, the diamond particles roll over the lap surface just as with other grinding surfaces, but they soon become embedded and cut the surface, producing chips.

Polishing

Polishing is the final step in producing a deformation free surface that is flat, scratch free, and mirror-like in appearance. Such a surface is necessary for subsequent metallographic interpretation, both qualitative and quantitative. The polishing technique used should not introduce extraneous structures, such as disturbed metal, pitting, dragging out of inclusion, "comet tailing," and staining. Polishing is usually conducted in several stages. Rough polishing is generally done with 6 or 3 μm diamond abrasive charged onto napless or low nap cloths. Hard materials, such as through hardened steels and cemented carbides, may require an additional

polishing step. For such materials, initial rough polishing may be followed by polishing with 1 μm diamond on a napless, low nap, or medium nap cloth. A compatible lubricant should be used sparingly to prevent overheating or deformation of the surface. Intermediate polishing should be performed thoroughly so that final polishing may be of minimal duration. Manual or hand polishing is usually conducted using a rotating *wheel*.

Mechanical Polishing

The term *mechanical polishing* is frequently used to describe the various polishing procedures involving the use of fine abrasives on cloth. The cloth may be attached to a rotating wheel or a vibrating bowl. The specimens are held by hand, held mechanically, or merely confined within the polishing area.

Hand Polishing. Aside from the use of improved polishing cloths and abrasives, hand polishing techniques still follow the basic practice established many years ago:

- *Specimen Movement:* The specimen is held with one or both hands, depending on the operator's preference, and is rotated in a direction counter to the rotation of the polishing wheel. In addition, the specimen is continuously moved back and forth between the center and the edge of the wheel, thereby ensuring even distribution of the abrasive and uniform wear of the polishing cloth. Some metallographers use a small wrist rotation while moving the specimen from the center to the edge of one side of the wheel. The main reason for rotating the specimen is to prevent formation of "comet tails." This polishing artifact (Fig. 3) is a result of directional polishing of materials containing inclusions, fine precipitates, voids, or other similar features.
- *Polishing Pressure:* The correct amount of applied pressure must be determined by experience. In general, firm hand pressure is applied to the specimen in the initial movement.
- *Washing and Drying:* The specimen is washed and swabbed in warm running water, rinsed with ethanol, and dried in a stream of warm air. Scrubbing with cotton soaked with an aqueous soap solution followed by rinsing with water is also commonly employed. Alcohol usually can be used for washing when the abrasive carrier is not soluble in water or if the specimen cannot tolerate water.
- *Cleanness:* The precautions for cleanness must be strictly observed

For routine metallographic work, a fine diamond abrasive (1 μm) may be used as the last step. Traditionally, aqueous fine alumina slurries have been used for final polishing with medium nap cloths. Alpha alumina (0.3 μm) and gamma alumina (0.05 μm) slurries (or suspensions) are popular for final polishing, either in sequence or singularly. Colloidal silica (basic pH about 9.5) and acidic alumina suspensions are newer final polishing

Fig. 3 Comet tails due to directional polishing and pull out of hard particles. Original magnification 200×.Source: Ref 1

abrasives being used for difficult to prepare materials. Vibratory polishers are often used for final polishing, particularly with more difficult to prepare materials, for image analysis studies, or for publication quality work.

Automatic Polishing. Mechanical polishing can be automated to a high degree using a wide variety of devices ranging from relatively simple systems to rather sophisticated minicomputer or microprocessor controlled devices. Units also vary in capacity from a single specimen to a half dozen or more at a time. Most units can be used for all grinding and polishing steps. These devices enable the operator to prepare a large number of specimens per day, often with a higher degree of quality than that of hand polishing and at reduced consumable costs. Automatic polishing devices also are desirable for preparing radioactive specimens by remote control or for using corrosive attack polishing procedures safely without hand contact.

Polishing Cloths. The requirements of a good polishing cloth include the ability to hold an abrasive; long life; absence of any foreign material that may cause scratches; and absence of any processing chemical such as dye or sizing that may react with the specimen. More than a hundred cloths of different fabrics (woven or nonwoven) with a wide variety of naps (or napless) are available for metallographic polishing. Napless or low nap cloths are recommended for rough polishing using diamond abrasive compounds. Low, medium, and occasionally high nap cloths are used for final polishing, but this step should be as brief as possible to minimize relief.

Polishing Abrasives. Polishing usually involves the use of one or more of the following abrasives: diamond, aluminum oxide (Al_2O_3), magnesium oxide (MgO), and/or silicon dioxide (SiO_2). For certain materials, cerium oxide, chromium oxide, or iron oxide may be used. With the exception of diamond, these abrasives normally are used in a distilled water suspension. If the metal to be polished is not compatible with water, other suspensions, such as ethylene glycol, alcohol, kerosene, or glycerol, may be required. The diamond abrasive should be extended only with the carrier recommended by the manufacturer.

Electrolytic Polishing

Even with the most careful mechanical polishing, some disturbed metal, however small the amount, will remain after preparation of a metallographic specimen. This remainder is no problem if the specimen is to be etched for structural investigation, because etching is usually sufficient to remove the slight layer of disturbed metal. If the specimen is to be examined in the as-polished condition using polarized light or if no surface disturbance can be tolerated, either electrolytic polishing (also called *electropolishing*) or chemical polishing is preferred. Alternatively, vibratory polishing with (basic) colloidal silica, acidic alumina suspensions, or attack polishing agents added to these abrasives (or to α or γ alumina suspensions) will remove minor amounts of residual damage providing good polarized light response. A simple laboratory setup (Fig. 4) is sufficient for most electropolishing requirements, and the more sophisticated commercial units are all based on the same principle. Direct current from an external source is applied to the electrolytic cell under specific conditions, and anodic dissolution produces leveling and brightening of the specimen surface.

Not all materials respond equally well to electrolytic polishing. Wrought solid solution alloys, such as aluminum, nickel, nickel-iron, and titanium

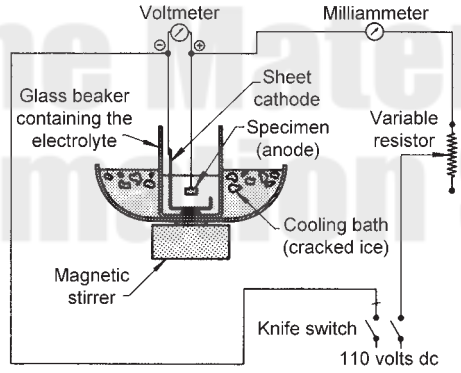


Fig. 4 Simple laboratory system for electropolishing and electroetching. Source: Ref 1

alloys, are particularly good candidates for electrolytic polishing. Electropolishing is usually reserved for single phase alloys, because second phases and inclusions may be preferentially attacked during polishing.

Chemical Polishing

Chemical polishing involves simple immersion of a metal specimen into a suitable solution to obtain a metallographic polish. The results of chemical polishing are similar to those of electropolishing. They vary from an etched specimen surface that has been macrosmoothed but not brightened to a bright dipped surface that has been brightened but not macrosmoothed.

Etching

Metallographic etching encompasses all processes used to reveal particular structural characteristics of a metal that are not evident in the as-polished condition. Examination of a properly polished specimen before etching may reveal structural aspects, such as porosity, cracks, and nonmetallic inclusions. In certain nonferrous alloys, grain size can be revealed adequately in the as-polished condition using polarized light.

Electrolytic Etching. The procedure for electrolytic etching is basically the same as for electropolishing, except that voltage and current densities are considerably lower. The specimen is connected to be the anode, and some relatively insoluble but conductive material, such as stainless steel, graphite, or platinum, is used for the cathode. Direct current electrolysis is used for most electrolytic etching, and for small specimens (13 by 13 mm, or ½ by ½ in., surface to be etched), one or two standard 1½ V flashlight batteries provide an adequate power source. A setup like the one shown in Fig. 4 is usually all that is required.

Etching for Microstructure. In this chapter, microscopic examination is limited to a maximum magnification of 2500×—the approximate useful limit of light microscopy. Microscopic examination of a properly prepared specimen will clearly reveal structural characteristics, such as grain size, segregation, and the shape, size, and distribution of the phases and inclusions, that are present. The microstructure revealed also indicates prior mechanical and thermal treatment that the metal has received.

Etching is done by immersion or by swabbing (or electrolytically) with a suitable chemical solution that basically produces selective corrosion. Swabbing is preferred for those metals and alloys that form a tenacious oxide surface layer with atmospheric exposure, such as stainless steels, aluminum, nickel, niobium, and titanium and their alloys. It is best to use surgical grade cotton that will not scratch the polished surface. Etch time varies with etch strength and can only be determined by experience. In general, for high magnification examination, the etch should be shallower, while for low magnification examination a deeper etch yields better image

contrast. Some etchants produce selective results in that only one phase will be attacked or colored. A vast number of etchants have been developed; the more commonly used etchants are given in Tables 1 to 6.

Microscopic Examination

Metallurgical microscopes differ from biological microscopes primarily in the manner by which the specimen is illuminated. Unlike biological

Table 1 Typical etchants used for microscopic examination

General reagents for irons and steels (carbon, low, and medium-alloy steels)

Etching reagent	Composition	Remarks	Uses
Nital	2 mL HNO_3, 90 mL ethyl or methyl alcohol (95% or absolute; also amyl alcohol)	Not as good as picral for high-resolution work with heat-treated structures; excellent for outlining ferrite grain boundaries; etching time, a few seconds to 1 min	For carbon steels, gives maximum contrast between pearlite and a ferrite or cementite network; reveals ferrite boundaries; differentiates ferrite from martensite
Picral	4 g picric acid, 100 mL ethyl or methyl alcohol (95% or absolute; use absolute alcohol only when acid contains 10% or more moisture), 4 or 5 drops zephiran chloride (17%)—wetting agent	Not as good as nital for revealing ferrite grain boundaries; gives superior resolution with fine pearlite, martensite, tempered martensite, and bainitic structures; detects carbides; etching time, a few seconds to 1 min or more	For all grades of carbon steels, annealed, normalized, quenched, quenched and tempered, spheroidized, austempered
Sodium metabisulfite	A 8 g $Na_2S_2O_5$, 100 mL distilled water	General reagent for steel; results similar to picral; etching time, a few seconds to 1 min	Darkens as-quenched martensite
	B 1 g $Na_2S_2O_5$, dilute to 100 mL with distilled water	Immerse specimen in the solution for 2 min or until the polished surface turns a bluish-red; do not mount specimen in a steel clamp	Tint etches lath-type or plate-type martensite in Fe-C alloys
Vilella's reagent	5 mL HCl, 1 g picric acid, 100 mL ethyl or methyl alcohol (95% or absolute)	Best results obtained when martensite is tempered	For revealing austenitic grain size in quenched and quenched and tempered steels
Heat tinting	Heat only	Heat by placing specimen face up on a hot plate that has been preheated to 205–370 °C (400–700 °F); time and temperature both have decided effects; bath of sand or molten metal may be used	Pearlite first to pass through a given color, followed by ferrite; cementite less affected, iron phosphide still less
Heat etching	Heat only	Specimen is heated 10–60 min at 815–1205 °C (1500–2200 °F) in carefully purified hydrogen, and must have no contact with scale or reducible oxides; after etching, specimen is cooled in mercury to avoid oxidation	For revealing austenitic grain size of polished specimens
Klemm's reagent	50 mL saturated (in H_2O) $Na_2S_2O_3$ solution, 1 g $K_2S_2O_5$	Etching time, 40–120s, ferrite appears black-brown, while carbides, nitrides, and phosphides remain white; also, phosphorus distribution can be detected more sensitively than with usual phosphorus reagents based on copper salts	Tint etches pearlite, hardened structures of unalloyed steel, and cast iron

Source: Ref 2

Table 2 Typical etchants used for microscopic examination

General reagents for alloy steels (high alloy, stainless, and tool steels)

Etching reagent	Composition	Remarks	Uses
Ferric chloride and hydrochloric acid	5 g $FeCl_3$, 50 mL HCl, 100 mL distilled water	Immerse until structure is revealed	Reveals structure of austenitic nickel and stainless steels
Mixed acids in glycerin	A 10 mL HNO_3, 20 mL HCl, 30 mL glycerin	Mix HCl and glycerin thoroughly before adding HNO_3; before etching, heat specimen in hot water; best results are obtained with alternate polishing and etching; use hood; do not store; action can be modified by varying the proportion of HCl	Etches structures of Fe-Cr alloys, high-speed steels, austenitic steels, and manganese steels; for austenitic alloys
	B 10 mL HNO_3, 20 mL HCl, 20 mL glycerin, 10 mL H_2O_2		Reveals the structures of Cr-Ni and Cr-Mn steels, and of all Fe-Cr austenitic alloys
Cupric chloride and hydrochloric acid	5 g $CuCl_2$, 100 mL HCl, 100 mL ethyl alcohol, 100 mL distilled water	Use cold	For austenitic and ferritic steels, the ferrite being most easily attacked (carbides and austenite are not attacked)
Nitric and hydrofluoric acids	5 mL HNO_3, 1 mL HF (48%), 44 mL distilled water	Use cold for about 5 min under hood; HF is harmful to skin	For revealing general structure of austenitic stainless steel with avoidance of strain markings
Heat tinting	Heat only in air for 10–60 s at about 595–650 °C (1100–1200 °F)	Carbides remain white, and austenite darkens less rapidly than ferrite; specimens preferably etched first with a chemical reagent; use hood	For austenitic stainless steels containing ferrite and carbides
Ferric chloride and nitric acid	Saturated solution of $FeCl_3$ in HCl, to which a little HNO_3 is added	Use full strength under hood	Structure of stainless steels
Mixed acids in cupric chloride	30 mL HCl, 10 mL HNO_3, saturate with cupric chloride, and let stand 20–30 min before using	Apply by swabbing	For stainless alloys and others high in nickel or cobalt
Nitric and acetic acids	30 mL HNO_3, 20 mL CH_3COOH	Apply by swabbing under hood; do not store	For stainless alloys and others high in nickel or cobalt
Marble's reagent	4 g $CuSO_4$, 20 mL HCl, 20 mL distilled water	Immerse to reveal structure	Structure of stainless steels
Ferricyanide solution	50 g $K_3Fe(CN)_6$, 50 g KOH, 100 mL distilled water	Must be fresh; use boiling 2–5 min under hood; do not acidify; deadly HCN may be released	To distinguish between ferrite and sigma phase in Fe-Cr, Fe-Cr-Ni, Fe-Cr-Mn, and related alloys; colors sigma blue, ferrite yellow
Vilella's reagent	5 mL HCl, 1 mL picric acid, 100 mL ethyl or methyl alcohol (95% or absolute)	Immerse to reveal structure	Can etch numerous types of Fe-Cr, Fe-Cr-Ni, and Fe-Cr-Mn steels; also attacks the grain boundaries in Cr-Ni austenitic steels
Cupric sulfate and perchloric acid	10 g $CuSO_4$, 45 mL perchloric acid (70%), 55 mL distilled water	Boil 15 min; do not use in presence of organic materials; use hood; do not concentrate acid; highly explosive	Etches stainless steels, and shows chromium segregation by revealing areas poor in chromium
Acetic, nitric, and hydrochloric acids	25 mL CH_3COOH, 15 mL HNO_3, 15 mL HCl, 5 mL distilled water	Apply by swabbing under hood; do not store	Fe-Al alloys, general microstructure
Hydrochloric and chromic acids	25 mL HCl, 50 mL CrO_3 (10% chromic acid aqueous solution)	Activity is controlled by the amount of chromic acid; use hood	Suitable for heat-treated type 300 stainless steels
Hydrochloric acid in alcohol	50 mL HCl, 50 mL ethyl alcohol	More gradual etching can be obtained with less concentrated solutions (10–20%)	Suitable for etching steels containing chromium and nickel
Mixed acids in ethyl alcohol	2.5 g $FeCl_3$, 5 g picric acid, 2 mL HCl, 90 mL ethyl alcohol	Etching time, 15 s for austenitic cast irons to 1 h or more for high-chromium ferritic irons	For high-chromium, high-carbon cast irons
Nitric acid	5 to 10 mL HNO_3, 100 mL ethyl or ethyl alcohol (95% or better)	Use hood; HNO_3 and ethyl alcohol are a dangerous mixture above 5% HNO_3	General structure of high-speed tool steels
Sodium metabisulfite I	15 g $Na_2S_2O_5$, 100 mL distilled water	Etching time, a few seconds to 1 min	General structure of high-speed tool steels

(*continued*)

Source: Ref 2

Table 2 Continued

Etching reagent	Composition	Remarks	Uses
Hydrochloric and nitric acids	10 mL HCl, 3 mL HNO_3, 100 mL methyl alcohol	Etching time, 2–10 min	To reveal the grain size of quenched or quenched and tempered high-speed steel
Groesbeck's reagent	4 g $KMnO_4$, 4 g NaOH, 100 mL distilled water	Use at boiling point for 1–10 min	For high-speed and chromium or cobalt-rich alloys
Sodium metabisulfite II	Step No. 1: 25 mL HNO_3, 75 mL ethyl or methyl alcohol	Pre-etch 10 s; outlines grain boundaries and some structure; (Caution: HNO_3 and ethyl dangerous at this concentration)	Tint etchant for Fe-Ni alloys from 5–25% Ni; colors martensitic packets of different orientations different colors, and reveals the substructure of lath-type martensite
	Step No. 2: 15 to 35 g $Na_2S_2O_5$, dilute to 100 mL with distilled water	Immerse for 2 min or until polished surface turns bluish-red	Concentration of $Na_2S_2O_5$ varies with nickel content
Beraha's reagent I	3 g $K_2S_2O_5$, 10 g $Na_2S_2O_3$, dilute to 100 mL with distilled water	Pre-etch with 4% picral for 1–2 min; Immerse for 2 min or until polished surface turns bluish-red	Tint etchant for Fe-Mn alloys from 5–18% Mn; also good for revealing chemical and physical heterogeneity in Fe-C alloys; colors ferrite while cementite remains white in Fe-C alloys
Klemm's reagent	50 mL cold saturated (in H_2O) $Na_2S_2O_3$ solution, 5 g $K_2S_2O_5$	Distinguishes between gamma, epsilon, and alpha phases; epsilon-martensite remains white, alpha-martensite is colored black, and gamma, gray; contrast can be improved in chromium-rich steels by addition of glacial acetic acid	Tint etches Mn, Mn-C, and Mn-Cr steels
Beraha's reagent II	Stock solution A 1 vol. HCl (35%) + 5 vol. distilled water Stock solution B 1 vol. HCl (35%) + 2 vol. distilled water Stock solution C 1 vol. HCl (35%) + 1 vol. distilled water Stock solution D 20 g ammonium bifluoride dissolved in 1000 mL of stock solution A Stock solution E 40 g ammonium bifluoride dissolved in 1000 mL of stock solution B Stock solution F 50 g ammonium bifluoride dissolved in 1000 mL of stock solution C Stock solution G 10 to 15 g iron chloride dissolved in 1000 mL of stock solution B or C Stock solution H 10 g copper chloride dissolved in 1000 mL of stock solution B or C	Tint etchant No. 1: 100 mL of stock solution A or D plus 100 to 200 mg potassium metabisulfite immerse at room temperature and keep specimen moving until desired coloration is attained. Note: Containers and forceps of suitable plastic materials should be used for a reagent which contains ammonium bifluoride; exposure to skin is dangerous; use hood Tint etchant No. 2: Same as No. 1 except potassium metabisulfite 300–600 mg Tint etchant No. 3: 100 mL of stock solution (B, C, E, F, G or H) + 300–800 mg potassium metabisulfite; if coloration takes place without etching, lower amount of potassium metabisulfite	Tint etchant for martensitic stainless steel; colors the matrix only; carbides and nitrides are unaffected and in contrast with the colored matrix For ferritic and austenitic stainless steel. Colors the matrix only; carbides and nitrides are unaffected and in contrast with the colored matrix For corrosion and heat-resisting alloys. Colors the matrix only; carbides and nitrides are unaffected and in contrast with the colored matrix

Source: Ref 2

Table 3 Typical etchants used for microscopic examination

Miscellaneous reagents (segregation, depth of case, primary structure, and strain lines)

Etching reagent	Composition	Remarks	Uses
Stead's reagent	1 g $CuCl_2$, 4 g $MgCl_2$, 1 mL HCl, 100 mL alcohol (absolute)	Dissolve salts in least possible quantity of hot water. Etch for about 1 min, repeating if necessary	To show segregation of phosphorus or other elements in solid solution; copper tends to deposit first on areas lowest in phosphorus; structure may be more clearly delineated by light hand polish to remove the copper deposit after etching

(*continued*)

Source: Ref 2

Table 3 Continued

Etching reagent	Composition	Remarks	Uses
Fry's reagent	5 g $CuCl_2$, 40 mL HCl, 30 mL distilled water, 25 mL ethyl alcohol	May be used cold; etching time, about 10 s	To reveal strain lines
Oberhoffer's reagent	30 g $FeCl_3$, 1 g $CuCl_2$, 0.5 g $SnCl_2$, 50 mL HCl, 500 mL ethyl alcohol, 500 mL distilled water	Immerse to reveal structure	For showing phosphorus segregation and dendritic structure
Alkaline chromate	16 g CrO_3, 145 mL distilled water, 80 g NaOH	Add NaOH slowly, and use when not over one day old, boiling at 120 ºC (250 ºF) for 7–20 min; use hood; this solution is highly caustic; should be prepared and stored in plastic	Shows oxygen segregation by darkening martensite rapidly, ferrite more slowly, and zones of high oxygen content much more slowly
Cupric sulfate and cupric chloride	1.25 g $CuSO_4$, 2.50 g $CuCl_2$, 10 g $MgCl_2$, 2 mL HCl, 100 mL distilled water; dilute above solution to 1000 mL with 95% ethyl alcohol	Proportions must be accurate; etch by immersion to avoid confusing edge effects; etching time, 30 s–1 min	For showing total depth of case, structure, and various zones of nitrided Cr-V steels and Nitralloy
Picric and nitric acids	10 parts picric acid (4%), 1 part HNO_3 (4%)	Best results are obtained when specimen is annealed in lead at 800 ºC (1475 ºF) before etching	For depth of case and structure of Nitralloy
Nital	1 mL HNO_3, 100 mL ethyl or methyl alcohol (95% or absolute)		For structure and depth of case of nitrided steels
Marble's reagent	4 g $CuSO_4$, 20 mL HCl, 20 mL distilled water		Total depth of nitrided case
Ammonium acetate	10 mL CH_3COONH_4, 100 mL distilled water		Stains high sulfur areas in steel, and stains lead particles brown in leaded steels

Source: Ref 2

Table 4 Typical etchants used for microscopic examination

Reagents for nonmetallic inclusions and intermetallic compounds

Etching reagent	Composition	Remarks	Uses
Ferricyanide solution	1 to 4 g $K_3Fe(CN)_6$, 10 g KOH, 100 mL distilled water	Must be freshly made; etch 15 min in boiling solution; 7 g of NaOH may be substituted for 10 g of KOH; use hood; do not acidify as HCN may be released	Differentiates between carbides and nitrides; cementite is blackened, pearlite turns brown, and massive nitrides remain unchanged
Murakami's reagent	10 g $K_3Fe(CN)_6$, 10 g KOH, 100 mL distilled water	May be used cold, but preferably hot; should be freshly made; 7 g of NaOH may be substituted for 10 g KOH. Etching time 5–10 min; use hood; do not acidify as HCN may be released	Alloy, high-speed, and tungsten steels; colors various carbides differently; cementite not affected
Strong ferricyanide solution	40 g $K_3Fe(CN)_6$, 100 g KOH, 100 mL distilled water	Use fresh, boiling for 15 min; use hood; do not acidify as HCN may be released	Stainless and alloy steels; differentiates between carbides and nitrides
Chromic acid and heat tinting	8 g CrO_3, 100 mL distilled water	Etch first in 4% picric acid then for 1 min in chromic acid; heat tint by heating face-up on a hot plate at about 500 ºF for 1 min	Distinguishes between iron phosphide and cementite in phosphide eutectic of cast iron; iron phosphide is colored darker
Sodium picrate, alkaline	2 g picric acid, 25 g NaOH, 100 mL distilled water	Use boiling, 5–10 min; do not boil dry	Colors cementite, but not carbides high in chromium; attacks sulfides; delineates grain boundaries in hypereutectoid steels in slowly cooled condition
Hydrogen peroxide and sodium hydroxide	10 mL H_2O_2 (30%), 20 mL NaOH 10% solution in distilled water	Must be fresh; etching time, 10–20 min; highly caustic solution	Attacks and darkens iron tungstide in carbon-free Fe-W alloys; when carbon is present, this solution darkens the compound (FeW, WC) in proportion to the amount of carbide present; tungsten carbide is darkened
Sodium hydroxide and potassium permanganate	4 g NaOH, 4 g $KMnO_4$, 100 mL distilled water	Use boiling; etching time, 1–10 min	Alloy and high-speed steels; colors precipitated carbides in manganese or chromium steels; differentiates between carbides and tungstides; vanadium carbide unattacked
Sodium hydroxide and bromine	20 g NaOH, 4 mL bromine, 80 mL distilled water	Does not keep; use fresh under hood	Colors iron phosphide
Sodium hydroxide and lead nitrate	1 part 50% NaOH, 2 parts 10% $Pb(NO_3)_2$	Use fresh, cold or boiling; highly caustic solution	Steels, colors cementite, attacks phosphides and silicates

Source: Ref 2

Table 5 Typical etchants used for microscopic examination

Reagents for macroscopic examination

Etching reagent	Composition	Remarks	Uses
Hydrochloric acid	50 mL HCl, 50 mL H_2O	Use at 70–80 ºC (160–180 ºF) for 1–60 min, depending on the size of sample, type of steel, and type of structure to be developed; use hood	Shows segregation, porosity, cracks depth of hardened zone in tool steel, and so on; may produce cracks in strained steel
Mixed acids	38 mL HCl, 12 mL H_2SO_4, 50 mL H_20	Use hot or boiling, 15–45 min or cold for 2–4 h; use hood	Steel, general macro; one of the best; shows segregation, cracks, hardened zone, soft spots, weld structures
Nitric acid in water	A 25 mL HNO_3, 75 mL H_2O	Use cold on large surfaces such as split ingots which cannot conveniently be heated	Same as HCl reagent
	B 0.5 to 1 mL HNO_3, 99.5 to 99 mL H_2O	Immerse 30–60 s after grinding specimen on 240-grit emery belt and thorough cleaning	To show weld structures
Nital	5 mL HNO_3, 95 mL ethyl alcohol	Etch 5 min followed by 1 s in 10% HCl in H_2O	Shows cleanliness, depth of hardening, carburized or decarburized surface, and so on
Ammonium persulfate	10 mL $(NH_4)_2S_2O_8$, 90 mL H_2	Surface should be rubbed with absorbent cotton during etching	Brings out grain structure, excessive grain growth, recrystallization at welds, flow lines in Nitralloy, and so on
Ammonium persulfate with potassium iodide, and so on	A 2.5 g $(NH_4)_2S_2O_8$, 100 mL H_2O	After grinding on No. 320 abrasive paper, swab 15 min with solution A, then 10 min with B, then 5 min with C, and 5 min with D, finally washing with water and drying with alcohol	Shows dendritic macrostructure of cast iron
	B Same as A, plus 1.5 g KI		
	C Same as B, plus 1.5 g $HgCl_2$		
	D Same as C, plus 15 mL H_2SO_4		
Stead's reagent	2.5 g $CuCl_2$, 10 g $MgCl_2$, 5 mL HCl, up to 250 mL ethyl alcohol	Salts are dissolved in HCl with the addition of the least possible quantity of water	Brings out phosphorus-rich areas and phosphorus banding; may be used for general segregation
Fry's reagent	A 90 g $CuCl_2$, 120 mL HCl, 100 mL H_2O	Most useful for low-carbon steels, particularly bessemer and other high-nitrogen grades. Before etching, sample should be heated to 150–250 ºC (300–480 ºF) for 5–30 min; depending on condition of steel; during etching, surface should be rubbed with cloth soaked in etching solution; wash in alcohol or rinse in HCl (1:1) after etching to prevent deposition of copper	Shows up strain lines due to cold work
	B 45 g $CuCl_2$, 180 mL HCl, 100 mL H_2O	Specimen can be washed in water without depositing copper; gives contrast	Shows strain lines due to cold work
Humfrey's reagent	120 g $Cu(NH_3)_4Cl_2$, 50 mL HCl, 1000 mL H_2O	Slight abrasion of surface after etching is recommended	Develops dendritic segregation
Kalling's reagent	1.5 g $CuCl_2$, 33 mL HCl, 33 mL H_2O, 33 mL alcohol	Etching time very short	Develops dendritic pattern in steel, attacks ferritic and martensitic stainless steels; ferrite darkened, martensite darker, austenite light
Marble's reagent	10 g $CuSO_4$, 50 mL HCl, 50 mL H_2O	May be used hot	Ni-Cr-Fe alloys, manganese and Cr-Mn steels, nitrided case; carbide precipitation in austenitic alloys
Vilella's reagent	1 g picric acid, 5 mL HCl, 100 mL alcohol	Use hot	Fe-Cr-Ni and Fe-Cr-Mn steels; reveals austenitic grain boundaries

Source: Ref 2

Table 6 Typical etchants used for microscopic examination

Electrolytes for polishing and etching

Etching reagent	Composition	Remarks	Uses
Chromic acid	10 g CrO_3, 100 mL H_2O	Specimen is used as anode; stainless steel or platinum as cathode, ¾ to 1 in. apart; 6 V usually used; etching time, 30–90 s	For various structures except the grain boundaries of ferrite; attack cementite very rapidly, austenite less rapidly, ferrite and iron phosphide very slowly if at all
Nitric acid in water	50 mL HNO_3, 50 mL H_2O	Room temperature; stainless steel cathode; 1.5 V for 2 min or more; use hood	For austenitic or ferritic stainless steels; reveals grain boundaries

(*continued*)

Source: Ref 2

Table 6 Continued

Etching reagent	Composition	Remarks	Uses
Hydrochloric acid in alcohol	10 mL HCl, 90 mL anhydrous ethyl alcohol	10–30 s at 6 V	Reveals delta ferrite, and the general structure of chromium and Cr-Ni steels
Sulfuric acid in water	5 mL H_2SO_4, 95 mL H_2O	Room temperature; stainless steel cathode; 6 V (0.1–0.5 amp), 5–15 s; use hood	For Fe-Cr-Ni alloys
Mixed acids in alcohol	45 mL lactic acid, 10 mL HCl, 45 mL ethyl alcohol	10–30 s at 6 V	For chromium steels (4 to 30% Cr), or for delta ferrite in austenitic stainless steels
Oxalic acid in water	10 mL oxalic acid, 100 mL H_2O	5–20 s at 6 V using a platinum or stainless steel cathode; gap between electrodes, ¾–1 in.	For austenitic stainless steels and high-nickel alloys; distinguishes between sigma phase and carbides; sigma phase is attacked first, then carbides; ferrite and austenite can be attacked slightly. To investigate the carbides, operate at 1.5–3 V for a longer time
Sodium hydroxide in water	40 g NaOH, 100 mL H_2O (add slowly)	60 s at 1–3 V; highly caustic solution	Reveals the sigma phase; colors successively sigma phase, ferrite, and lastly carbides after a longer etching time
Potassium hydroxide in water	56 g KOH, 100 mL H_2O (add slowly)	60 s at 1–3 V; highly caustic solution	Same as sodium hydroxide in water, but sigma phase and ferrite are revealed simultaneously
Sodium cyanide in water	10 g NaCN, 100 mL H_2O	5 min or more at 6 V (not less than 5 V); use hood; if acidified, HCN develops	Colors carbides without altering austenite or grain boundaries
Ammonium persulfate in water	10 g $(NH_4)_2S_2O_8$, 100 mL H_2O	Use fresh, 6 V for more than 15 s	Surface attack occurs on the ferrite grains in low-carbon steels; reveals the fine structures of nickel austenitic steels, and of transformer sheet
Ammonium nitrate	Saturated aqueous solution of NH_4NO_3	Use a current density of 1 amp/sq cm	Detects overheating and burning; in overheating, this etchant leaves the boundaries of the pre-existing austenite grains white, while it blackens them in burned steel
Cadmium acetate in water	10 g cadmium acetate, 100 mL H_2O	2–5 V for 3–20 s	Brings out carbide grain boundaries
Mixed acids in water	5 g ammonium molybdate, 7.5 mL HNO_3, 10 mL HCl, 100 mL H_2O	12 V for 2–3 min	Good for type 300 stainless steels
Mixed acids in water	90 mL H_2PO_3, 8 mL HNO_3, 2 mL H_2O	Use cold	For polishing and etching Fe-Al alloys (to 16% Al)
Chrome regia	25 mL HCl, 5 to 50 mL CrO_3 solution (10%) in H_2O	Dilute with 2 parts alcohol and 2 parts glycerin; etch for 20–60 s at 6 V	Heat-treated type 300 stainless steels
Perchloric-acetic	10 parts glacial acetic acid, 1 part perchloric acid	2 min or more at 20–22 V and 0.1 amp/sq cm; temperature < 20 °C (70 °F); stainless steel cathode, maintain constant stirring of solution during polishing; use hood	Polish iron and steel. Note: Add perchloric acid very slowly to glacial acetic acid at < 15 °C (60 °F) and stir constantly to prevent localized heating; wash in cold running water and rinse in ethyl alcohol
Chrome-acetic	775 mL glacial acetic acid, 150 g Na_2CrO_4, 75 g CrO_3	2 min or more at 40–45 V; stir solution and keep below 20 °C (68 °F)	Very good for polishing iron and steel; slower but less dangerous than above etchant; polishing operation does not need to be constantly monitored

Source: Ref 2

microscopes, metallurgical microscopes must use reflected light. A simplified ray diagram for a metallurgical microscope is shown in Fig. 5, while a typical inverted metallurgical microscope (*metallograph*) is shown in Fig. 6. The prepared specimen is placed on the stage with the surface perpendicular to the optical axis of the microscope and is illuminated through the objective lens by light from the source. The light is focused by the condenser lens into a beam that is made approximately parallel to the optical axis of the microscope by the half-silvered mirror. The light is then reflected from the surface of the specimen through the objective, the half-silvered mirror, and the eyepiece to the observer's eye.

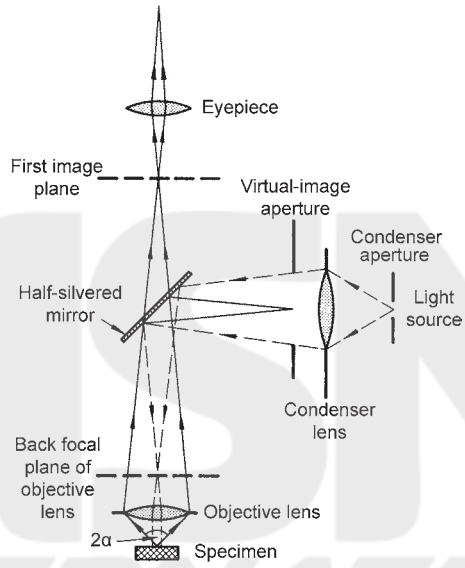


Fig. 5 Image formation in a metallurgical microscope employing bright-field illumination. Source: Ref 1

Fig. 6 Example of an inverted metallurgical reflecting microscope for photomicroscopy (referred to as a metallograph). Courtesy of Nikon Inc. Source: Ref 1

Light Sources

The amount of light lost during passage from the source through a reflecting type of microscope is appreciable because of the intricate path the light follows. For this reason, it is generally preferable that the intensity of the source be high, especially for photomicroscopy. Several light sources are used, including tungsten filament lamps, tungsten-halogen lamps, quartz-halogen lamps, and xenon arc bulbs.

Tungsten filament lamps generally operate at low voltage and high current. They are widely used for visual examination because of their low cost and ease of operation.

Tungsten-halogen lamps are the most popular light source due to their high light intensity. They produce good color micrographs when tungsten-corrected films are employed.

Xenon arc lamps produce extremely high intensity, and their uniform spectra and daylight color temperature makes them suitable for color photomicrography. The first xenon lamps produced ozone, but modern units have overcome this problem. Light output is constant and can only be reduced using neutral density filters.

Microscopic Techniques

Most microscopic studies of metals are made using bright-field illumination. In addition to this type of illumination, several special techniques (oblique illumination, dark-field illumination, opaque stop microscopy, phase contrast microscopy, and polarized light microscopy) have particular applications for metallographic studies.

Köhler Illumination. Most microscopes using reflected or transmitted light use Köhler illumination, because it provides the most intense, even illumination possible with standard light sources. The reflected light microscope has two adjustable diaphragms, the aperture diaphragm and the field diaphragm, located between the lamp housing and the objective. Both diaphragms are adjusted to improve illumination and the image. To obtain Köhler illumination, the image of the field diaphragm must be brought into focus on the specimen plane. This situation normally occurs automatically when the microstructural image is brought into focus. The filament image must also be focused on the aperture diaphragm plane. This focus produces uniform illumination of the specimen imaged at the intermediate image plane and magnified by the eyepiece.

In bright-field illumination, the surface of the specimen is normal to the optical axis of the microscope, and white light is used. Figure 5 shows the ray diagram for this type of illumination in a standard type of bench microscope. Light that passes through the objective and strikes a region of the specimen surface perpendicular to the beam will be reflected back up the objective through the eyepieces to the eyes, where it will appear to be bright or white. Light that strikes grain boundaries, phase boundaries, and other features not perpendicular to the optical axis will be scattered at an angle and will not be collected by the objective. These regions will appear to be dark or black in the image. Bright-field is the most common mode of illumination used by metallographers.

Oblique illumination reveals the surface relief of a metallographic specimen. This process involves offsetting the condenser lens system or, as is more usually done, moving the condenser aperture to a position

slightly off the optical axis. Although it should be possible to continually increase the contrast achieved by oblique illumination by moving the condenser farther and farther from the light axis, the numerical aperture of a lens is reduced when this happens because only a portion of the lens is used. For this reason, there is a practical limit to the amount of contrast that can be achieved. Illumination also becomes uneven as the degree of obliqueness increases. Because differential interference contrast systems have been available, oblique illumination is rarely offered as an option on new microscopes.

Dark-field illumination (also known as *dark-ground illumination*) often is used to distinguish features not in the plane of the polished and etched surface of a metallographic specimen. This type of illumination gives contrast completely reversed from that obtained with bright-field illumination: the features that are light in bright-field will be dark in dark-field, and those that are dark in bright-field will be light in dark-field. This highlighting of angled surfaces (namely, those of pits, crack, or etched grain boundaries) allows more positive identification of their nature than can be derived from a black image under bright-field illumination. Due to the high image contrast obtained and the brightness associated with features at an angle to the optical axis, it is often possible to see details not observed with bright-field illumination.

Polarized light microscopy is particularly useful in metallography, because many metals and metallic and nonmetallic phases are optically anisotropic. Polarized light is obtained by placing a polarizer in front of the condenser lens of the microscope and placing an analyzer before the eyepiece (Fig. 7). The polarizer produces plane polarized light that strikes

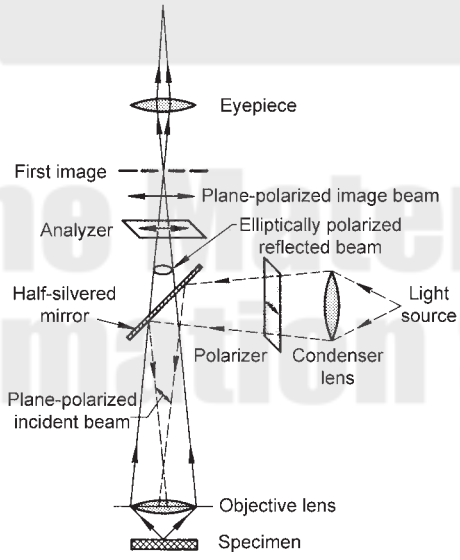


Fig. 7 Basic components of a polarizing light microscope. Source: Ref 1

the surface and is reflected through the analyzer to the eyepieces. If an anisotropic metal is examined with the analyzer set 90° to the polarizer, the grain structure will be visible. However, viewing of an isotropic metal (cubic metals) under such conditions will produce a dark, extinguished condition. Polarized light is particularly useful in metallography for revealing grain structure and twinning in anisotropic metals and alloys and for identifying anisotropic phases and inclusions.

Differential Interference Contrast Microscopy. When crossed polarized light is used along with a double quartz prism (Wollaston prism) placed between the objective and the vertical illuminator, two light beams are produced that exhibit coherent interference in the image plane. This occurrence leads to two slightly displaced (laterally) images differing in phase ($\lambda/2$) that produces height contrast. The image produced reveals topographic detail somewhat similar to that produced by oblique illumination but without the loss of resolution. Images can be viewed with natural colors similar to those observed in bright-field, or artificial coloring can be introduced by adding a sensitive tint plate.

Microphotography

All metallographs come equipped with one or more camera ports, as well as provisions for attaching charge-coupled device (CCD) cameras (or other types, although the CCD is by far the most common type used). While biologists frequently use 35 mm cameras to record microstructures, they are less popular with metallographers. A small percentage of metallographers still prefer to use wet processed sheet film, usually 4 by 5 in. (10 by 12.5 cm) in size. Orthochromatic film is no longer available in this size, and panchromatic films must be employed. These are less convenient to use because loading, unloading, and developing must be done in total darkness. Otherwise, results are the same. Contact printing is most commonly performed.

The majority of metallographers switched to instant films (Polaroid), which were introduced in the 1960s. At that time, few (if any) metallographs had exposure meters, and wastage was significant because instant films have no latitude (exposures must be exactly controlled to get good images, unlike wet processed films). Instant films are convenient because dark room work is avoided. However, except for the P/N type, there is no negative so extra prints cannot be made in the same way as by traditional photography. Instead, multiple photographs must be made anticipating future needs.

Electronic photography is now the dominant mode as it features all the convenience of instant photography with none of the disadvantages. In addition, digital images can be stored on a disk, eliminating the need for storing negatives. With the advent of high quality, high resolution printers,

publication ready photomicrographs can be produced. Finally, the Internet allows the metallographer to take a digital photomicrograph and immediately send it to other colleagues any where in the world.

Grain Size

Since the grain size of a metal or alloy has important effects on the structural properties (both strength and ductility), a number of methods have been developed to measure the grain size of a sample. In all methods, some form of microexamination is used in which a small sample is mounted, polished, and then etched to reveal the grain structure. The most direct method is then to count the number of grains present in a known area of the sample so the grain size can be expressed as the number of grains/area. The American Society for Testing Materials (ASTM) grain size index number is derived from the number of grains/in.^{2} when counted at a magnification of 100×. The ASTM index, N, is given by:

$$n = 2^{(N-1)}$$

where n is the number of grains/in.^{2} at 100× magnification. This can be rewritten as:

$$\log n = (N-1)\log 2$$

or

$$N = \frac{\log n}{0.3010} + 1$$

A listing of ASTM grain sizes is given in Table 7. Note that a larger ASTM grain size number indicates a finer or smaller grain size. The improvement in yield strength with finer grain sizes is illustrated in Fig. 8 for a number of metals.

Table 7 ASTM grain size, $n = 2^{N-1}$

	Average number of grains/unit area	
Grain size No. (N)	No./in.^{2} at 100× (n)	No./mm^2 at 1× (n)
1	1.00	15.50
2	2.00	31.00
3	4.00	62.00
4	8.00	124.00
5	16.00	496.00
6	32.00	992.00
7	64.00	1984.0
8	128.00	3968.0
9	256.00	3968.0
10	512.00	7936.0

Source: Ref 3

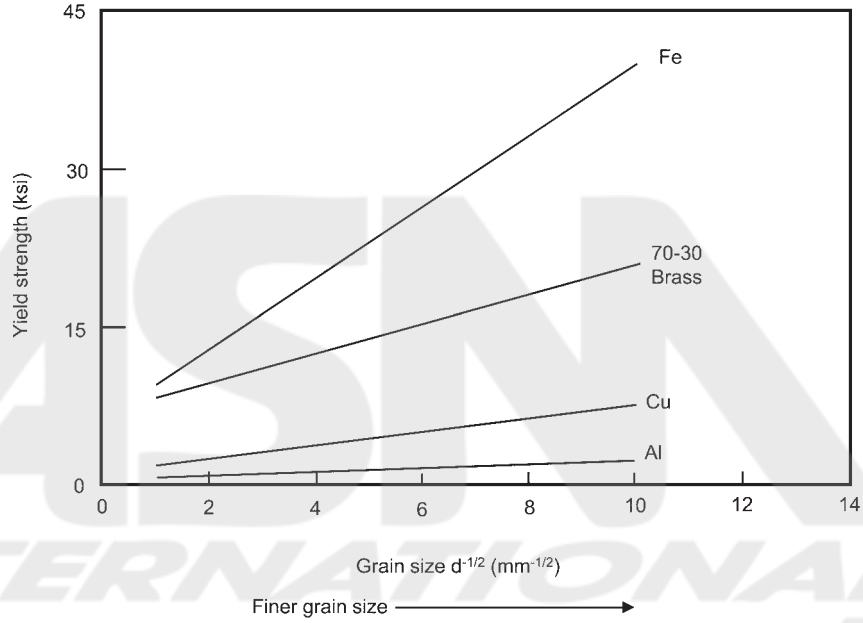


Fig. 8 Effect of grain size on yield strength. Source: Adapted from Ref 3

ACKNOWLEDGMENT

This chapter was adapted from Metallographic Practices Generally Applicable to All Metals by G.F. Vander Voort in *Metals Handbook Desk Edition*, Second Edition, 1998.

REFERENCES

1. G.F. Vander Voort, Metallographic Practices Generally Applicable to All Metals, *Metals Handbook Desk Edition*, 2nd ed., ASM International, 1998, p 1356–1371
2. R.C. Anderson, *Inspection of Metals: Destructive Testing*, ASM International, 1988
3. F.C. Campbell, *Elements of Metallurgy and Engineering Alloys*, ASM International, 2008

SELECTED REFERENCES

- *Metallography and Microstructures*, Vol 9, *ASM Handbook*, ASM International, 2004
- G.F. Vander Voort, *Metallography: Principles and Practice*, ASM International, 1984

CHAPTER 9

Liquid Penetrant, Magnetic Particle, and Eddy-Current Inspection

LIQUID PENETRANT, magnetic particle, and eddy current inspection are used to detect surface flaws. Magnetic particle and eddy current inspection can also detect subsurface flaws. Liquid penetrant and eddy current inspection can be used to inspect ferrous and nonferrous metals, while magnetic particle inspection is restricted to materials that can be magnetized. Both magnetic particle and eddy current inspection can be automated. In the case of automated eddy current inspection, quite high rates of production can be attained. While liquid penetrant inspection is primarily a manual operation, the equipment requirements are less than for magnetic particle or eddy current inspection.

Liquid Penetrant Inspection

Liquid penetrant inspection is a nondestructive method used to find discontinuities that are open to the surface of solid, essentially nonporous materials. Indications of flaws can be found regardless of the size, configuration, internal structure, and chemical composition of the workpiece being inspected, as well as flaw orientation. Liquid penetrants can seep into (and be drawn into) various types of minute surface openings (as fine as 0.1 μm, or 4 μin., in width) by capillary action. Therefore, the process is well suited to detect all types of surface cracks, laps, porosity, shrinkage areas, laminations, and similar discontinuities. It is used extensively to inspect ferrous and nonferrous metal wrought and cast products, powder metallurgy parts, ceramics, plastics, and glass objects.

The liquid penetrant inspection method is relatively simple to perform; there are few limitations due to specimen material and geometry, and it is

inexpensive. The equipment is very simple, and the inspection can be performed at many stages during part production, as well as after the part is placed in service. Relatively little specialized training is required. In some instances, liquid penetrant sensitivity is greater for ferromagnetic steels than that of magnetic particle inspection.

The major limitation of liquid penetrant inspection is that it can detect only imperfections that are open to the surface; some other method must be used to detect subsurface defects and discontinuities. Another factor that can inhibit the effectiveness of liquid penetrant inspection is the surface roughness of the object. Extremely rough and porous surfaces are likely to produce false indications.

Although liquid penetrant is often used to inspect some types of powder metallurgy parts, the process is generally not well suited to inspect low density powder metallurgy parts and other porous materials, because the penetrant enters the pores and registers each pore as a defect.

Physical Principles

Liquid penetrant inspection depends mainly on the ability of liquid penetrant to effectively wet the surface of a solid workpiece; flow over the surface to form a continuous, reasonably uniform coating; and migrate into cavities that are open to the surface. The cavities of interest usually are very small, often invisible to the unaided eye. The ability of a given liquid to flow over a surface and enter surface cavities depends principally on:

- Cleanness of the surface
- Configuration of the cavity
- Size of the cavity
- Surface tension of the liquid
- Ability of the liquid to wet the surface

The cohesive forces between molecules of a liquid cause surface tension. An example of the influence of surface tension is the tendency of a free liquid, such as a droplet of water, to contract into a sphere. In such a droplet, surface tension is counterbalanced by the internal hydrostatic pressure of the liquid. When the liquid comes into contact with a solid surface, the cohesive force responsible for surface tension competes with the adhesive force between the molecules of the liquid and the solid surface. These forces jointly determine the contact angle between the liquid and the surface. If the angle is less than 90°, the liquid is regarded as having good wetting ability.

Description of the Process

Regardless of the type of penetrant used and other variations in the basic process, liquid penetrant inspection requires at least five essential steps:

1. Surface Preparation. All surfaces of a workpiece must be thoroughly cleaned and completely dried before inspection. Discontinuities exposed to the surface must be free from oil, water, and other contaminants for at least 25 mm (1 in.) beyond the area being inspected to increase the probability of detection.

2. Penetrant Application. Liquid penetrant is applied in a suitable manner to form a film of the penetrant over the surface for at least 13 mm (0.5 in.) beyond the area being inspected. The penetrant is left on the surface for a sufficient time to allow penetration into flaws. Times are based on experience but generally range from 2 to 20 minutes.

3. Removal of Excess Penetrant. Uniform removal of excess penetrant is necessary for effective inspection, but over cleaning must be avoided. Penetrants can be washed off directly using water, treated first with an emulsifier and then rinsed with water, or removed using a solvent.

4. Developer Application. Developer can be applied by dusting (dry powdered) and immersion and spray (water developers) applications. Nonaqueous wet developers can only be applied by spraying. The developer should be allowed to dwell on the surface for a sufficient time (usually 10 minutes minimum) to permit it to draw penetrant out of any surface flaws to form visible indications of such flaws. Longer times could be necessary for tight cracks. The developer also provides a uniform background to assist visual inspection.

5. Inspection. After being sufficiently developed, the surface is visually examined for indications of penetrant bleed back from surface openings. This examination must be performed in a suitable inspection environment. Visible penetrant inspection is performed in good white light. When fluorescent penetrant is used, inspection is performed in a suitably darkened area using black (ultraviolet) light, which causes the penetrant to emit visible light. The actions of penetrant and developer are illustrated in Fig. 1.

Penetrant Systems

Liquid penetrant inspection applications have been developed to handle the wide variations in three basic penetrant systems. They are broadly classified as (a) the water-washable system, (b) the postemulsifiable system, and (c) the solvent-removable system.

The water-washable penetrant system is designed so that the penetrant is directly washable from the surface of the workpiece using water. It can be used to process workpieces quickly and efficiently. However, it is important that washing is carefully controlled, because water-washable penetrants are susceptible to over washing. The degree and speed of removal depend on processing conditions such as spray nozzle characteristics, water pressure and temperature, duration of rinse cycle, surface con-

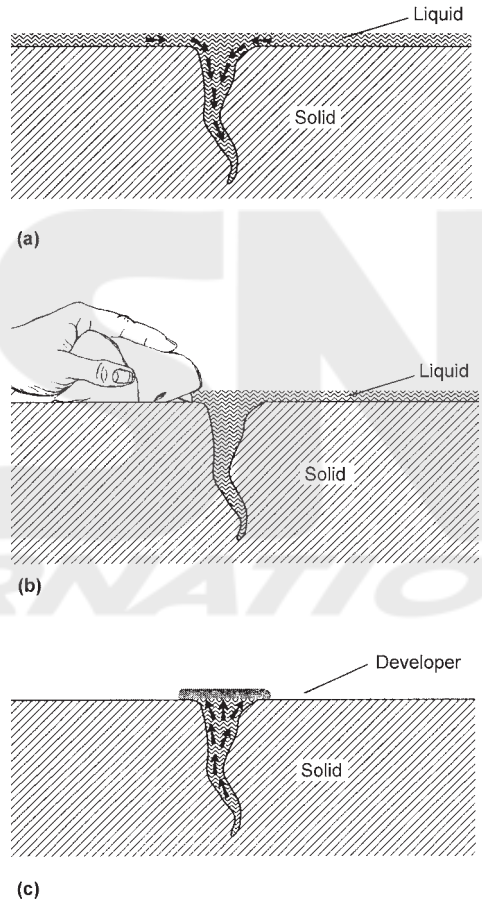


Fig. 1 Actions of penetrant and developer. (a) Penetrant liquid is drawn into an open crack by capillary action. (b) Excess surface penetrant is removed by wiping with a cloth, washing directly with water, treating with an emulsifier and rinsing, or removing with a solvent. (c) Developer applied to the surface draws out penetrant liquid that seeps into the developer forming a visible indication of the surface crack. A colored dye or a fluorescence compound is usually added to the penetrant liquid. Depending on the amount of penetrant that seeps into the developer, the crack width can appear 100 times larger than its actual size. Source: Ref 1

dition of the workpiece, and inherent removal characteristics of the penetrant employed.

The Postemulsifiable System. High sensitivity penetrants that are not water-washable are used to ensure detection of minute discoveries in some materials. Because they are not water-washable, the danger of washing the penetrant out of the flaws is reduced. These penetrants require an additional operation in the inspection process. An emulsifier must be applied after the application of penetrant and proper penetration (dwell) time. The emulsifier makes the penetrant soluble in water so the excess penetrant can be removed by water rinsing. Therefore, the emulsification time must

be carefully controlled so the surface penetrant becomes water soluble but penetrant in the flaws does not. Postemulsifiable penetrants include lipophilic (oil base) and hydrophilic (water base).

The Solvent-Removable System. Occasionally, it is necessary to inspect only a small area of a workpiece or to inspect a workpiece on site rather than at a regular inspection station. In such situations, solvent-removable penetrants are used. Normally, the same type of solvent is used both for precleaning and for removal of excess penetrant. This penetrant process is convenient and broadens the range of applications of penetrant inspection.

The solvent-removable penetrants have an oil base. Optimum solvent removal is accomplished by wiping off as much of the excess penetrant as possible with a paper towel or a lint free cloth, then slightly dampening a clean cloth with solvent and wiping off the remaining penetrant. Final wiping with a dry paper towel or clean cloth is required. The penetrant also can be removed by flooding the surface with solvent, in the same manner as for water-washable penetrants. The flooding technique is particularly useful for large workpieces, but it must be very carefully used to prevent removal of the penetrant from the flaws. The solvent-removable system is used mainly in special applications; it is not practical for production applications because it is labor intensive.

Liquid Penetrant Materials

Two basic types of liquid penetrants are fluorescent and visible. Each type is available for any one of the three systems (water-washable, postemulsifiable, or solvent-removable).

Fluorescent penetrant inspection uses penetrants that fluoresce brilliantly under ultraviolet light. The sensitivity of a fluorescent penetrant depends on its ability to form indications that appear as small sources of light in an otherwise dark area. Sensitivity levels of fluorescent penetrants are: ultra-low (Level ½), low (Level 1), medium (Level 3), high (Level 4), and ultrahigh (Level 5).

Visible penetrant inspection uses a penetrant that is usually red in color and produces vivid red indications in contrast to the light background of the applied developer under visible light. The visible penetrant indications must be viewed under adequate white light. The sensitivity of visible penetrants is regarded as Level 1 and adequate for many applications.

Penetrant selection and use depend on the criticality of the inspection, the condition of the workpiece surface, the type of processing, and the desired sensitivity.

Physical and Chemical Characteristics. Both fluorescent and visible penetrants, whether water-washable, postemulsifiable, or solvent-removable, must have certain chemical and physical characteristics to perform their intended functions. Principal requirements of penetrants are:

- Chemical stability and uniform physical consistency
- A flash point not lower than 95 °C (200 °F); penetrants that have lower flash points constitute a potential fire hazard
- A high degree of wettability
- Low viscosity to permit better coverage and minimum dragout
- Ability to penetrate discontinuities quickly and completely
- Sufficient brightness and permanence of color
- Chemical inertness with materials being inspected and with containers
- Low toxicity to protect personnel
- Slow drying characteristics
- Ease of removal
- Inoffensive odor
- Low cost
- Resistance to ultraviolet light and heat fade

Emulsifiers

Emulsifiers are liquids used to render excess oily penetrant on the workpiece surface water washable. Emulsifiers are oil base and water base.

Oil base emulsifiers function by diffusion. The emulsifier diffuses into the penetrant film and renders it spontaneously emulsifiable in water. The rate at which it diffuses into the penetrant establishes its emulsification time. Because the emulsifier is fast acting, the rinse operation should be done quickly to avoid over emulsification.

Water base emulsifiers usually are supplied as liquid concentrates that are diluted in water to concentrations of 5 to 30% for dip tank applications and of 0.05 to 5% for spray applications. Water base emulsifiers function by displacing excess surface penetrant from the surface of the part by detergent action. The force of the water spray or air agitation of open dip tanks provides the scrubbing action while the detergent displaces the excess surface penetrant.

Solvent Cleaners

Solvent cleaners differ from emulsifiers in that they remove excess surface penetrant through direct solvent action. The penetrant is dissolved by the solvent. Solvent cleaners are flammable and nonflammable. Flammable cleaners are free of halogens, but are potential fire hazards. Nonflammable cleaners usually contain halogenated solvents, which render them unsuitable for some applications, usually because of their high toxicity or because they have undesirable effects on some materials.

Developers

Because the amount of penetrant that emerges from a small surface opening is minute, the visible evidence of its presence must be enhanced. Developers are used to spread the penetrant available at the defect, thus

increasing the amount of light emitted, or the amount of contrast, that makes the defect visible to the unaided eye.

The properties/characteristics developers must have for optimal performance are:

- It must be adsorptive to maximize blotting
- It must have fine grain size and a particle shape that will disperse and expose the penetrant at a flaw to produce strong and sharply defined indications of flaws
- It must be capable of providing a contrast background for indications when color contrast penetrants are used
- It must be easy to apply
- It must form a thin, uniform coating over a surface
- It must be easily wetted by the penetrant at the flaw (the liquid must be allowed to spread over the particle surfaces)
- It must be nonfluorescent if used with fluorescent penetrants
- It must be easy to remove after inspection
- It must not contain ingredients harmful to parts being inspected or to equipment used in the inspection operation
- It must not contain ingredients harmful or toxic to the operator

Four forms of developers commonly used are dry powder (Form A), water soluble (Form B), water suspendible (Form C), and nonaqueous solvent suspendible (Form D).

Dry Developers. Dry powder developers are widely used with fluorescent penetrants, but should not be used with visible dye penetrants because they do not produce a satisfactory contrast coating on the surface of the workpiece. Ideally, dry powder developers should be light and fluffy to allow easy application and should cling to dry surfaces in a fine film. Powder adherence should not be excessive, because the amount of penetrant at fine flaws is insufficient to seep back into a thick coating.

Hand processing equipment usually includes a developer station, which usually is an open tank for dry developers. Workpieces are dipped into the powder, or powder is picked up with a scoop or with the hands, and dropped onto the workpiece. Excess powder is removed by shaking and tapping the workpiece. Some powders are so light and fluffy that parts are dipped into them as easily as into a liquid.

Other effective methods of application are rubber spray bulbs and air operated spray guns. An electrostatic charged powder gun that can apply an extremely even and adherent coating of dry powder on metal parts is also used. For simple application, especially when only a portion of the surface of a large part is being inspected, a very soft bristle brush is often adequate.

Powder can dry the skin and irritate the lining of air passages. Operators should use rubber gloves and respirators. Modern equipment often includes an exhaust system on the developer spray booth or on the devel-

oper dust chamber, which prevents dust from escaping. Powder recovery filters often are included in such installations.

Wet Developers. Three types of wet developers are: suspensions of developer powder in water (the most widely used), aqueous solutions of suitable salts, and suspensions of powder in volatile solvents.

Water suspendible developers can be used with both visible and fluorescent penetrants. With a fluorescent penetrant, the dried developer coating must not fluoresce or absorb or filter out black light used for inspection. Water suspendible developers permit high speed application of developer in mass inspection of small to medium size workpieces using the fluorescent method. A basket of small, irregularly shaped workpieces that has gone through the steps of penetrant application, penetrant dwell, and washing can be coated with developer in one quick dip in a water suspension. This method not only is quick, but also, it provides thorough, complete coverage of all surfaces of the pieces being inspected. No dry powder application method has all these advantages to the same degree.

Wet developer is applied just after excess penetrant is washed away and immediately before drying. After drying, surfaces are uniformly coated with a thin film of developer. Developing time is decreased because heat from the drier helps to bring penetrant back out of surface openings, and the developing action occurs immediately with the developer film already in place. Workpieces are ready for inspection in a shorter period of time, before excessive bleed out from large openings occurs, so better definition of flaw indications often is obtained.

Water suspendible developers are supplied as a dry powder concentrate, which is then dispersed in water in recommended proportions, usually from 0.04 to 0.12 kg/L (1/3 to 1 lb/gal). The amount of powder in suspension must be carefully maintained. Too much or too little developer on the surface of a workpiece can seriously affect sensitivity.

Water soluble developers can be used for both fluorescent and visible postemulsifiable and solvent-removable penetrants. Water soluble developers are not recommended for use with water-washable penetrants, because of the potential to wash the penetrant from within the flaw if the developer is not very carefully controlled. Water soluble developers are supplied as a dry powder concentrate, which is then dispersed in water in recommended proportions, usually from 0.12 to 0.24 kg/L (1 to 2 lb/gal). Advantages of this form of developer are:

- The prepared bath is completely soluble and does not require any agitation
- The developer is applied prior to drying, thus decreasing the developing time
- The dried developer film on the workpiece is completely water soluble and is easily and completely removed following inspection by simple water rinsing

Nonaqueous solvent suspendible developers are commonly used for both the fluorescent and the visible penetrant process. This form of developer produces a white coating on the surface of the part, which yields the maximum white color contrast with the red visible penetrant indication and extremely brilliant fluorescent indication.

Nonaqueous solvent suspendible developers are supplied in the ready to use condition and contain particles of developer suspended in a mixture of volatile solvents. The solvents are carefully selected for their compatibility with the penetrants. Nonaqueous solvent suspendible developers also contain surfactants in a dispersant that coat the particles and reduce their tendency to clump or agglomerate.

Nonaqueous solvent suspendible developers are the most sensitive form of developer used with fluorescent penetrants because the solvent action contributes to the absorption and adsorption mechanisms. In many cases where tight, small flaws occur, dry powder, water soluble, and water suspendible developers do not contact the entrapped penetrant. This results in the failure of the developer to create the necessary capillary action and surface tension that serve to pull the penetrant from the flaw. The nonaqueous solvent suspendible developer enters the flaw and dissolves into the penetrant. This action increases the volume and reduces the viscosity of the penetrant.

The manufacturer must carefully select and compound the solvent mixture. There are two types of solvent based developers: nonflammable (chlorinated solvents) and flammable (nonchlorinated solvents). Both types are widely used. Selection is based on the nature of the application and the type of alloy being inspected.

Solvent developers are sometimes applied with a paintbrush, but this is likely to result in smeared indications; application by a pressure spray can is a preferred method.

Selection of Developer. Because developers play such an important role in penetrant inspection, it is very important to select the appropriate developer for a given job. For example, on very smooth or polished surfaces, dry powder does not adhere satisfactorily, and wet developers do a better job. Conversely, on very rough surfaces dry powder is far more effective.

Some general rules regarding developer selection are:

- Use a wet developer instead of a dry developer on very smooth surfaces
- Use a dry developer versus a wet developer on very rough surfaces
- Wet developers are better suited for high production inspection of small workpieces because of their greater ease and speed of application
- Wet developers cannot be used reliably where sharp fillets unavoidably accumulate developer, which can mask flaw indications

- Solvent developers are effective for revealing fine, deep cracks, but are not satisfactory for finding wide, shallow flaws
- Cleaning and reinspecting a rough surface is difficult if a wet developer was used for a prior inspection

The developer does not produce indications but simply absorbs the penetrant already present in or at the flaw and makes it more visible.

Equipment Requirements

With the exception of a source of ultraviolet (black) light for use with fluorescent penetrants, there is no special equipment that is absolutely essential for liquid penetrant inspection. Reasonably effective inspection can be performed with a minimum of simple and relatively basic equipment. However, this approach should be considered only when: (a) no more than a few workpieces are involved, (b) specific portions of very large workpieces are being inspected, (c) maximum sensitivity is not required, or (d) inspection must be performed in the field. Therefore, most liquid penetrant inspection is done with equipment designed specifically for the purpose.

A variety of equipment is available. "Package units" that incorporate all the necessary stations and controls are widely used, especially where relatively small workpieces in a variety of sizes and shapes are being inspected. A typical package unit for inspection using a water-washable, fluorescent penetrant system is shown in Fig. 2. This system is designed to process a steady flow of workpieces, which move through seven stations: application of penetrant, draining excess penetrant, water rinsing, inspection under ultraviolet light to check thoroughness of rinsing, drying, ap-

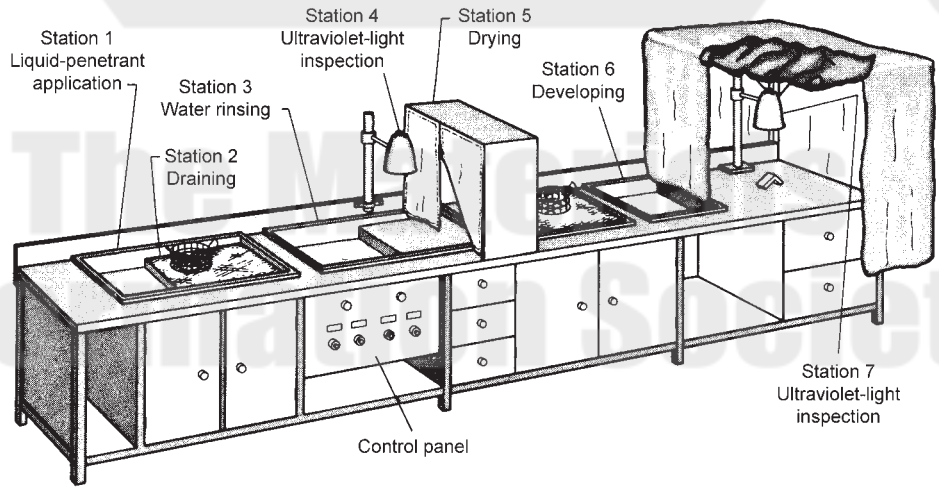


Fig. 2 Typical seven station package equipment unit for inspecting workpieces using a water washable, fluorescent penetrant system. Source: Ref 1

plication of developer, and final ultraviolet-light inspection for flaws. The unit does not include stations for preliminary cleaning and post cleaning; these operations often are performed in another area. The equipment shown in Fig. 2 is available in a wide range of sizes and can be modified in many ways to fit specific needs. For example, if a postemulsifiable system is used, workpieces are coated with emulsifier after the penetrant has been allowed to drain and prior to rinsing.

Precleaning. Regardless of the penetrant chosen, adequate precleaning of workpieces prior to penetrant inspection is absolutely necessary for accurate results. Inadequate removal of surface contamination can result in missed relevant indications because:

- The penetrant does not enter the flaw
- The penetrant loses its ability to reveal the flaw because it reacts with a substance contained in the flaw
- The surface immediately surrounding the flaw retains too much penetrant, which masks the true appearance of the flaw

Also, nonrelevant (false) indications can be caused by residual materials holding penetrants.

Cleaning methods generally are classified as chemical, mechanical, solvent, and combinations of these.

Chemical cleaning methods include alkaline or acid cleaning, pickling or chemical etching, and molten salt bath cleaning.

Mechanical cleaning methods include tumbling, wet blasting, dry abrasive blasting, wire brushing, and high pressure water or steam cleaning. Mechanical cleaning can mask flaws by smearing adjacent metal over them and by filling them with abrasive material.

Solvent cleaning methods include vapor degreasing, solvent spraying, solvent wiping, and ultrasonic immersion using solvents. Ultrasonic immersion is by far the most effective means of ensuring clean parts, but it can be a very expensive capital equipment investment.

Cleaning methods and their common uses are listed in Table 1. Major factors in the selection of a cleaning method are the type of contaminant to be removed, the type of alloy being cleaned, and the chemical composition of the workpiece being cleaned. It is good practice to test the method on known flaws to ensure that it will not mask the flaws.

The surface finish of the workpiece must always be considered. When further processing is scheduled, such as machining or final polishing, or when a surface finish of 3.20 μm (125 μin.) or coarser is allowed, an abrasive cleaning method is frequently a good choice. Generally, chemical cleaning methods have fewer degrading effects on surface finish than mechanical methods (unless the chemical used is strongly corrosive to the material being cleaned). Steam cleaning and solvent cleaning rarely have any effect on surface finish.

Table 1 Applications of various methods of precleaning for liquid penetrant inspection

Method	Use
Mechanical	
Abrasive tumbling	Removing light scale, burrs, welding flux, braze stopoff, rust, casting mold, and core material; should not be used on soft metals such as aluminum, magnesium, and titanium
Dry abrasive grit blasting	Removing light and heavy scale, flux, stopoff, rust, casting mold and core material, sprayed coatings, and carbon deposits—in general, any friable deposit. Can be fixed or portable
Wet abrasive grit blasting	Same as dry except, where deposits are light, better surface and better control of dimensions are required
Wire brushing	Removing light deposits of scale, flux, and stopoff
High-pressure water and steam	Typically used with an alkaline cleaner or detergent; removing typical machine-shop soils such as cutting oils, polishing compounds, grease, chips, and deposits from electrical discharge machining; used when surface finish must be maintained; inexpensive
Ultrasonic cleaning	Typically used with detergent and water or with a solvent; removing adherent shop soil from large quantities of small parts
Chemical	
Alkaline cleaning	Removing braze stopoff, rust, scale, oils, greases, polishing material, and carbon deposits; typically used on large articles where hand methods are too labor intensive; also used on aluminum for gross metal removal
Acid cleaning	Strong solutions for removing heavy scale; solutions for light scale; weak (etching) solutions for removing lightly smeared metal
Molten salt bath cleaning	Conditioning and removing heavy scale
Solvent	
Vapor degreasing	Removing typical shop soil, oil, and grease; usually uses chlorinated solvents; not suitable for titanium
Solvent wiping	Same as for vapor degreasing except a hand operation; can use nonchlorinated solvents; used for localized low-volume cleaning

Source: Ref 1

The choice of a cleaning method can be dictated by Occupational Safety and Health Administration and Environmental Protection Agency health and safety regulations. Factors to consider include quantities of materials that will be used, toxicity, filtering, neutralization and disposal techniques, and worker safety.

Penetrant Station. The principal requirement of a penetrant station is to provide a means to coat workpieces with penetrant–the entire surface for small workpieces, or over small areas of large workpieces when only local inspection is required. The station also should provide a means to drain excess penetrant back into the penetrant reservoir, unless the expendable technique is being used. Draining racks usually serve the additional purpose of providing a storage place for parts during the time required for penetration (dwell time).

Emulsifier Station. Emulsifier liquid is contained in a tank large enough to permit immersion of the workpieces, either individually or in batches. Accessory equipment includes covers to reduce evaporation and drain valves for cleanout when the bath has to be renewed. Suitable drain racks are also a part of this station, to permit excess emulsifier to drain back into the tank.

Large workpieces must be coated with emulsifier as fast as possible. Multiple spraying or copious flowing of emulsifier from troughs or perforated pipes can be used on some types of automatic equipment. For local coating of large workpieces, spraying often is satisfactory, using the expendable technique described for application of penetrant.

Rinse Station. Water rinsing (washing) of small workpieces frequently is done by hand, either individually or in batches in wire baskets. The workpieces are held in the wash tank and cleaned with a hand held spray using water at tap pressure and temperature (water temperature should not, however, be below 10 °C, or 50 °F).

Drying Station. The recirculating hot air drier is one of the most important equipment components. The drier must be large enough to easily handle the type and number of workpieces being inspected. Heat input, air flow, rate of movement of workpieces through the drier, and temperature control are all factors that must be balanced. The drier may be of the cabinet type illustrated in Fig. 2, or it can be designed so that the workpieces pass through on a conveyor. If conveyor operation is used, the speed must be coordinated with the required drying cycle.

Developer Station. The type and location of a developer station depends on whether dry or wet developer is used. For dry developer, the developer station is downstream from the drier, whereas for wet developer, it immediately precedes the drier, following the rinse station.

The dry developer station usually consists of a simple bin containing the powder. Dried workpieces are dipped into the powder and the excess powder is shaken off. For larger workpieces that are difficult to immerse in the powder, a scoop can be used to throw powder over the surfaces, after which the excess is shaken off. The developer bin should be equipped with an easily removable cover to protect the developer from dust and dirt when not in use. Dust control systems are sometimes needed when dry developer is used.

The inspection station essentially is a worktable on which workpieces can be handled under proper lighting. For fluorescent methods, the table usually is surrounded by a curtain or hood to exclude most of the white light from the area. For visible penetrants, a hood is not necessary.

Black (ultraviolet) lights can consist of batteries of 100 or 400 watt lamps for area lighting, or, in small stations, can be one or two 100 watt spot lamps mounted on brackets from which they can be lifted and moved about by hand. Because of the heat given off by black lights, good air circulation is essential in black light booths. For automatic inspection, workpieces are moved through booths equipped with split curtains, either by hand or by conveyor.

Postcleaning Station. Post inspection cleaning is often necessary to remove all traces of penetrant and developer. Drastic chemical or mechanical methods are seldom required for postcleaning. When justified by the

volume of work, an emulsion cleaning line is effective and reasonable in cost. In special circumstances, ultrasonic cleaning may be the only satisfactory way of cleaning deep crevices or small holes. However, solvents or detergent aided steam or water is almost always sufficient. The use of steam with detergent is probably the most effective of all methods.

Selection of Penetrant System

Size, shape, and weight of workpieces, as well as number of similar workpieces to be inspected, can influence the selection of a penetrant system.

Sensitivity and Cost. The required level of sensitivity and cost usually are the most important factors in selecting a system. The methods capable of the greatest sensitivity are also the most costly. There are many inspection operations that require the ultimate in sensitivity, but there are also many where extreme sensitivity is not required, but where extreme sensitivity can also produce misleading results.

On a practical basis, the three major penetrant systems are broken down into six systems or variations of systems. The six systems, in order of decreasing sensitivity and decreasing cost are:

- Postemulsifiable fluorescent
- Solvent-removable fluorescent
- Water-washable fluorescent
- Postemulsifiable visible
- Solvent-removable visible
- Water-washable visible

Table 2 compares the sensitivities and uses of the six systems.

Table 2 Comparison of penetrant systems

Water washable	Postemulsifiable	Solvent removable
Visible dye penetrants		
Lowest in sensitivity	Higher sensitivity than water washables	Where water rinse is not feasible, or desirable
Suited for large surface areas	Suited for large surface areas Suited for large quantities of similar objects	For spot inspections Recommended for small areas and simple geometries
Fluorescent penetrants		
Lowest in sensitivity of fluorescent penetrants	Higher sensitivity than water washable fluorescent penetrants	Higher sensitivity than solvent-removable visible penetrant
Suited for large surface areas	Suited for large quantities of similar articles	Where water rinse is not feasible or desirable
Suited for large quantities of similar objects	Suited for wide, shallow discontinuities and tight cracks	For spot inspections
Suited for deep, narrow discontinuities	Contaminants must be removed prior to inspection	Recommended for small areas and simple geometries
Recommended for rough surfaces (i.e., sand castings)	Suited for stress, intergranular, and grinding cracks	

Note: For environmental reasons, water washable penetrant systems are used even though the solvent system would be preferable.
Source: Ref 1

Magnetic Particle Inspection

Magnetic particle inspection is used to locate surface and subsurface discontinuities in ferromagnetic materials. The method is based on the fact that when a material or part being tested is magnetized, discontinuities that lie in a direction generally transverse to the direction of the magnetic field cause a leakage field to form at and above the part surface. The presence of the leakage field, and, therefore, the presence of the discontinuity, is detected by the use of finely divided ferromagnetic particles applied over the surface. Some of the particles are gathered and held by the leakage field. The magnetically held particles form an outline of the discontinuity and generally indicate its location, size, shape, and extent. Magnetic particles are applied over a surface either as dry particles or as wet particles in a liquid carrier such as water and oil. Nonferromagnetic materials cannot be inspected by this method. Such materials include aluminum alloys, magnesium alloys, copper and copper alloys, lead, titanium and titanium alloys, and austenitic stainless steels.

The principal industrial uses of magnetic particle inspection are final inspection; receiving inspection; in-process inspection and quality control; maintenance and overhaul in the transportation industries; plant and machinery maintenance; and inspection of large components. Although in-process magnetic particle inspection is used to detect discontinuities and imperfections in material and parts as early as possible in the sequence of operations, final inspection is required to ensure that rejectable discontinuities and imperfections detrimental to part use and function have not developed during processing.

Advantages. The magnetic particle method is a sensitive means to locate small, shallow surface cracks in ferromagnetic materials. Cracks large enough to be seen by the naked eye can produce an indication, but very wide cracks will not produce a particle pattern if the surface opening is too wide for the particles to bridge.

Discontinuities that do not actually break through the surface also are indicated in many instances using magnetic particle inspection within certain limitations. Fine, sharp discontinuities close to the surface (a long stringer of nonmetallic inclusions, for example) can produce an indication. Indications of deeper discontinuities are less distinct.

Limitations. The operator must be aware of certain limitations of magnetic particle inspection. For example, thin coatings of paint and nonmagnetic coverings, such as plating, adversely affect sensitivity. Other limitations are:

- Workpiece material must be ferromagnetic
- The direction of the magnetic field must intercept the principal plane of the discontinuity at right angles for best results. This could require two or more sequential inspections with different magnetizations

- Demagnetization following inspection is often necessary
- Post cleaning to remove remnants of magnetic particles and carrier solutions on the surface could be required after testing and demagnetization
- Inspection of very large parts could require a very large current
- Local heating and burning of finished parts and surfaces at the points of electrical contact is possible if care is not exercised
- Experience and skill in interpreting the significance of magnetic particle indications is necessary

Description of Magnetic Fields

Magnetized Ring. When a magnetic material is placed across the poles of a horseshoe magnet having square ends (forming a closed or ring like assembly), the magnetic lines of force flow from the north pole through the magnetic material to the south pole (see Fig. 3a). Magnetic lines of force flow preferentially through magnetic material rather than through nonmagnetic material or air. The magnetic lines of force are enclosed within the ring like assembly because no external poles exist, and iron filings or magnetic particles dusted over the assembly are not attracted to the magnet even though there are lines of magnetic force flowing through it.

If one end of the magnet is not square, leaving an air gap between the magnet end and the magnetic material, the poles still attract magnetic materials. Magnetic particles cling to the poles and bridge the gap between them, as shown in Fig. 3(b). A radial crack in a round magnetized piece creates north and south magnetic poles at the edges of the crack. Magnetic particles are attracted to the poles created by the crack, forming an indication of the discontinuity.

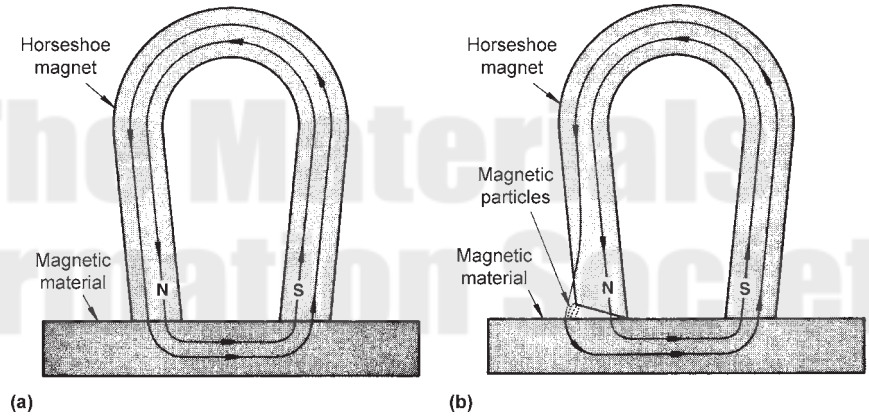


Fig. 3 (a) Horseshoe magnet with a bar of magnetic material across poles forms a closed, ring like assembly, which will not attract magnetic particles. (b) Ring like magnet assembly with an air gap, to which magnetic particles are attracted. Source: Ref 2

The magnetic fields at cracks and other physical and magnetic discontinuities in the surface are called *leakage fields*. The strength of a leakage field determines the number of magnetic particles that will gather to form indications; strong indications are formed at strong fields, and vice versa. The density of the magnetic field determines its strength and is partly governed by the shape, size, and material of the part being inspected.

Magnetized Bar. A straight piece of magnetized material (bar magnet) has a pole at each end. Magnetic lines of force flow through the bar from the south pole to the north pole. Because the magnetic lines of force within the bar magnet run the length of the bar, it is said to be longitudinally magnetized or to contain a longitudinal field.

If a bar magnet is broken into two pieces, a leakage field with north and south poles is created between the pieces, as shown in Fig. 4(a). The field exists even if the fracture surfaces are brought together (see Fig. 4b). If the magnet is cracked but not broken completely into two pieces, a similar result occurs. A north and a south pole form at opposite edges of the crack, just as though the break were complete (see Fig. 4c). It is this field that attracts the iron particles that outline the crack. The strength of the poles is different from that of the completely broken pieces; it is a function of the crack depth and the width of the air gap at the surface.

Circular Magnetization. Electric current passing through any straight conductor, such as a wire or bar, creates a circular magnetic field around

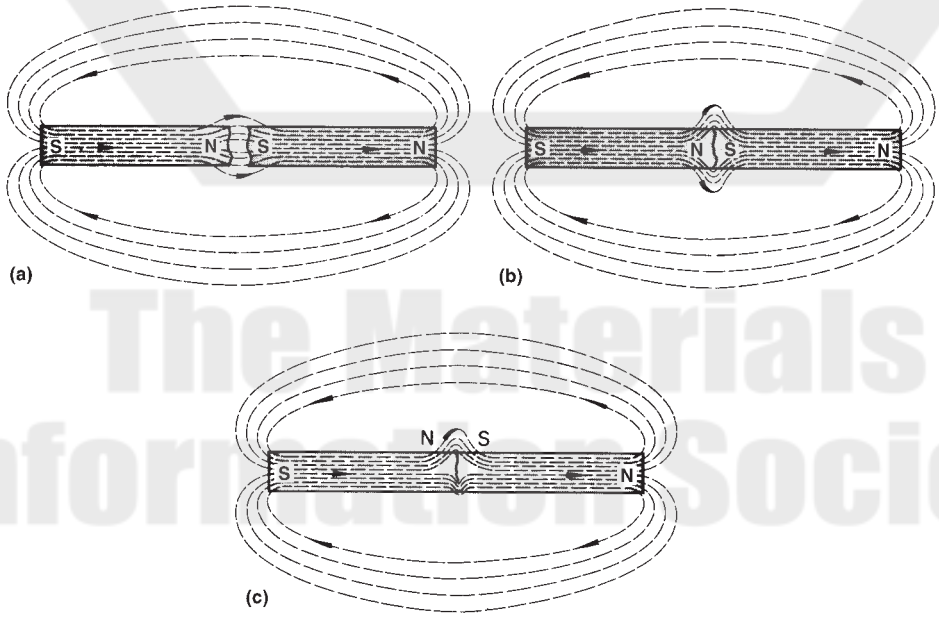


Fig. 4 Leakage fields between two pieces of a broken bar magnet (a) with magnet pieces apart, and (b) with magnet pieces together (simulating a flaw). (c) Leakage field at a crack in a bar magnet. Source: Ref 2

the conductor. The passage of current through a ferromagnetic conductor induces a magnetic field in both the conductor and surrounding space. A part magnetized in this manner is said to have a circular field or to be circularly magnetized, as shown in Fig. 5(a).

Longitudinal Magnetization. Electric current also can be used to create a longitudinal magnetic field in magnetic materials. When current is passed through a coil of one or more turns, a magnetic field is established lengthwise, or longitudinally, within the coil, as shown in Fig. 5(b). The nature and direction of the field around the conductor that forms the turns of the coil produces longitudinal magnetization.

Effect of Flux Direction. To form an indication, the magnetic field must approach a discontinuity at a sufficiently large angle to cause the magnetic lines of force to leave the part and return after bridging the discontinuity. An intersection approaching 90° produces the best results. For this reason, discontinuity direction, size, and shape are important. The direction of the magnetic field and the strength of the field in the area of the discontinuity also are important for optimum results.

Figure 6a illustrates a condition where the current is passed through the part, causing formation of a circular field around the part. Under normal circumstances, there would be no indication of the presence of a discontinuity such as one designated A in Fig. 6(a) because it is regular in shape and lies in a direction parallel to that of the magnetic field. A discontinuity having an irregular shape and predominantly parallel to the magnetic field, B, has a good chance to form a weak indication. Where the predominant direction of the discontinuity is at a 45° angle to the magnetic field, such

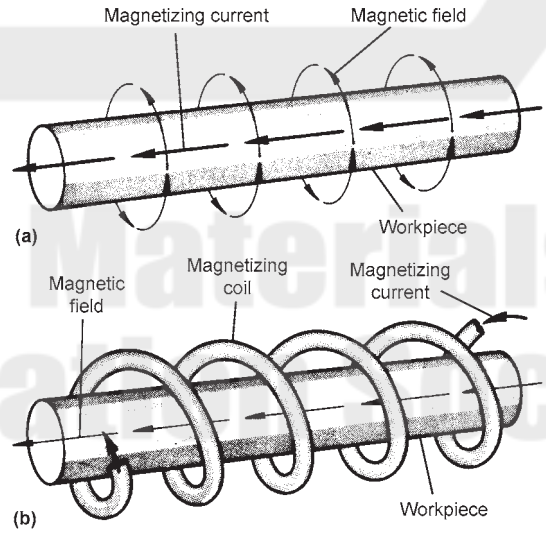


Fig. 5 Magnetized bars showing directions of magnetic field: (a) Circular. (b) Longitudinal. Source: Ref 2

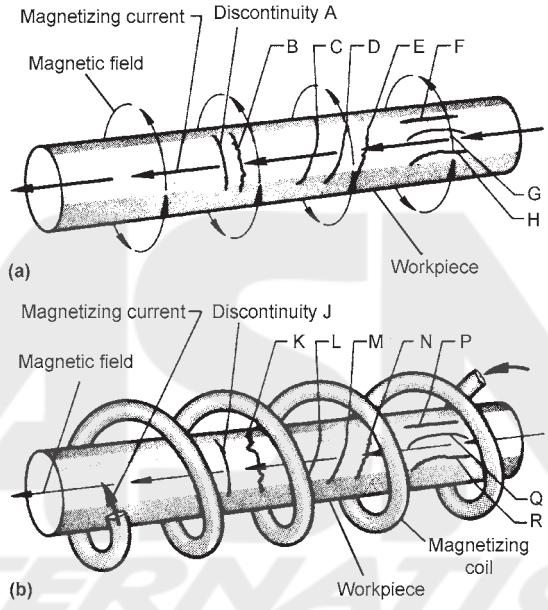


Fig. 6 Effect of direction of magnetic field or flux flow on detectability of discontinuities having various orientations. See text for discussion. (a) Circular magnetization. (b) Longitudinal magnetization. Source: Ref 2

as at C, D, and E, the conditions are more favorable for detection regardless of the shape of the discontinuity. Discontinuities whose predominant directions, regardless of shape, are at a 90° angle to the magnetic field (F, G, and H) produce the most pronounced indications.

Figure 6(b) shows a bar that has been longitudinally magnetized. Discontinuities L, M, and N, which are at about 45° to the magnetic field, would produce detectable indications, as they would in a circular field. Discontinuities J and K would display pronounced indications, but discontinuities P, Q, and R would probably not be detected.

Magnetization Methods. In magnetic particle inspection, the magnetic particles can be applied to the part while the magnetizing current is flowing and after the current has ceased to flow, depending largely on the magnetization retention (residual magnetism) of the part. The former technique is known as the *continuous method*; the latter, as the *residual method*.

If residual magnetism does not provide a leakage field strong enough to produce readable indications when magnetic particles are applied to the surface, the part must be continuously magnetized during application of particles. Consequently, the residual method can be used only on materials having sufficient retentivity; harder materials usually have higher retentivity. The continuous method is the only method used on low carbon steels and iron having little or no retentivity.

Magnetizing Current

Both direct and alternating currents are suitable to magnetize parts for magnetic particle inspection. Magnetic field strength, direction, and distribution are greatly affected by the current type used for magnetization.

The fields produced by direct and alternating current have different characteristics. The important difference in magnetic particle inspection is that fields produced by direct current generally penetrate the cross-section of the part, whereas fields produced by alternating current are confined to the metal at or near the surface of the part, which commonly is known as the skin effect. Therefore, alternating current should not be used to search for subsurface discontinuities.

Direct Current. The best source of direct current is rectified alternating current. Both single-phase and three-phase alternating currents are furnished commercially. Rectifiers convert reversing alternating current to unidirectional current. Rectified three-phase alternating current is nearly equivalent to straight direct current for purposes of magnetic particle inspection. The only difference between rectified three-phase alternating current and straight direct current is a slight ripple in the value of the rectified current, amounting for only about five percent of the maximum current value.

Alternating current, which must be single-phase when used directly to magnetize a part, is used directly from commercial power lines at a frequency of 50 or 60 Hz. When used for magnetizing, line voltage is stepped down by means of a transformer to the lower voltages required. Magnetizing currents of several thousand amperes often are used at the low voltages.

One problem in using alternating current is that residual magnetism in the part might be lower than that of the magnetism generated by the peak current of the alternating current cycle. This is because the level of residual magnetism depends on where in the cycle the current is discontinued.

Power Sources

Portable equipment is available in lightweight (16 to 41 kg, or 35 to 90 lb) power source units that are easily transported to the inspection site. Generally, portable units are designed to use 115, 230, and 460 V alternating current, and supply 750 to 1500 A magnetizing current outputs in half-wave or alternating current. Small, lightweight pulsed dc units that can produce up to 7000 A of magnetizing current also are available.

Mobile units generally are mounted on wheels to facilitate transport to an inspection site. Mobile equipment usually supplies full-wave, half-wave, and alternating magnetizing current outputs. Part inspection is accomplished by use of flexible cables, yokes, prod contacts, contact clamps, and coils. Instruments and controls are mounted on the front of the unit. Magnetizing current is usually controlled using a remote control switch

connected to the unit by an electrical cord. Quick coupling connectors for connecting magnetizing cables are on the front of the unit. Mobile equipment is usually powered by single phase or three phase, 60 Hz alternating current (230 and 460 V), and has an output range of 1500 to 10,000 A.

Stationary equipment is available in both general purpose and special purpose units. The general purpose unit is primarily used in the wet method, and has a built-in tank containing a bath pump, which continuously agitates the bath and forces the fluid through hoses onto the part being inspected. Pneumatically operated contact heads, together with a rigid type coil, provide capabilities for both circular and longitudinal magnetization. Self-contained ac and dc power supplies are available in amperage ratings from 1000 to 10,000 A.

Stationary power packs function as high-amperage, magnetizing current sources used in conjunction with special fixtures, and with cable wrap and clamp and contact techniques. Rated output varies from a customary 4000 to 6000 A to as high as 30,000 A. Higher amperage units are used for overall magnetization of large forgings and castings, which otherwise would require systematic prod inspection at much lower current levels. Some units feature three output circuits, which are systematically energized in rapid sequence, either electrically or mechanically, to effectively magnetize a part in several directions at virtually the same time. This allows exposure of discontinuities lying in any direction after only a single processing step.

Special purpose stationary units are designed to handle and inspect large quantities of similar items. Generally, conveyors, automatic markers, and alarm systems are included in such units to expedite part handling and disposition.

Methods of Generating Magnetic Fields

A basic requirement of magnetic particle inspection is to properly magnetize the part so leakage fields created by discontinuities attract magnetic particles. While permanent magnets can accomplish this to some degree, magnetization generally is produced using electromagnets and the magnetic field associated with the flow of electric current. Basically, magnetization is derived from the circular magnetic field generated when an electric current flows through a conductor. The direction of the field is dependent on the direction of current flow.

Yokes. Two basic types of yokes commonly used for magnetizing purposes are permanent magnets and electromagnetic yokes. Both are hand held and mobile.

Electromagnetic yokes (Fig. 7) consist of a coil wound around a U-shape core of soft iron. The legs of the yoke are either fixed or adjustable. Adjustable legs permit changing the contact spacing and the relative angle of contact to accommodate irregular shape parts. Unlike a permanent

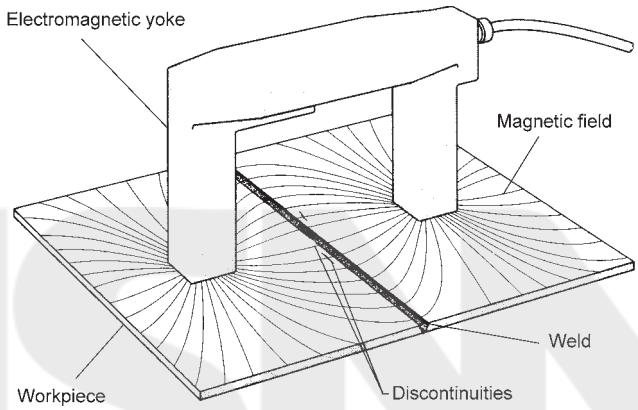


Fig. 7 Electromagnetic yoke, showing position and magnetic field to detect discontinuities parallel to a weld bead. Discontinuities across a weld bead can be detected by placing the contact surfaces of the yoke next to and on either side of the bead (rotating yoke about 90° from position shown here). Source: Ref 2

magnet yoke, an electromagnetic yoke can readily be switched on or off, a feature that makes it convenient to apply and remove the yoke from the test piece.

Electromagnetic yoke design is based on the use of either direct or alternating current, or both. The flux density of the magnetic field produced by the direct current type can be changed by varying the amount of current in the coil. The direct current type of yoke has greater penetration, whereas the alternating current type concentrates the magnetic field at the surface of the test piece, providing good sensitivity for revealing surface discontinuities over a relatively narrow area. Yokes using alternating current for magnetization have various applications and can be used for demagnetization as well. Discontinuities generally need to be centrally located in the area between pole pieces and oriented perpendicular to an imaginary line connecting them to be exposed.

In operation, the part completes the magnetic path for the flow of magnetic flux. The yoke is a source of magnetic flux, and the part becomes the preferential path completing the magnetic circuit between the poles. (In Fig. 7, only those portions of the flux lines near the poles are shown.)

Coils. Single loop and multiple loop coils (conductors) are used for longitudinal magnetization of components (see Fig. 5b and 6b). The field within the coil has a definite direction, corresponding to the direction of the lines of force running through it. The flux density passing through the interior of the coil is proportional to the product of the current I in amperes, and the number of turns in the coil N. Thus, the magnetizing force of such a coil can be varied by varying either the current or the number of turns in the coil. Coils for large parts can be made by winding several turns of a flexible cable around the part. Care must be taken to ensure that no indications are concealed beneath the cable.

Portable magnetizing coils are available that can be plugged into an electrical outlet. These coils can be used for in-place inspection of shaft like parts in railroad shops, aircraft maintenance shops, and shops for automobile, truck, and tractor repair. Transverse cracks in spindles and shafts are easily detected using such coils.

Most coils used for magnetizing are short, especially those wound on fixed frames. The relation of the length of the part being inspected to the width of the coil must be considered. For a simple part, the effective overall distance that can be inspected is 150 to 225 mm (6 to 9 in.) on either side of the coil. Thus, a part 300 to 450 mm (12 to 18 in.) long can be inspected using a normal coil approximately 25 mm (1 in.) thick. In testing longer parts, either the part must be moved at regular intervals through the coil, or the coil must be moved along the part.

The ease with which a part can be longitudinally magnetized in a coil is significantly related to the length-to-diameter (*L*/*D*) ratio of the part. This is due to the demagnetizing effect of the magnetic poles that are set up at the ends of the part. This demagnetizing effect is pronounced for *L*/*D* ratios of less than 10 to 1, and very significant for ratios of less than 3 to 1. Where the *L*/*D* ratio is extremely unfavorable, pole pieces of similar cross-sectional area can be introduced to effectively increase the length of the part and consequently improve the *L*/*D* ratio.

Central Conductors. For many tubular and ring shaped parts, it is advantageous to use a separate conductor to carry the magnetizing current, rather than the part itself. Such a conductor, commonly referred to as a central conductor, is threaded through the inside of the part (Fig. 8) and is

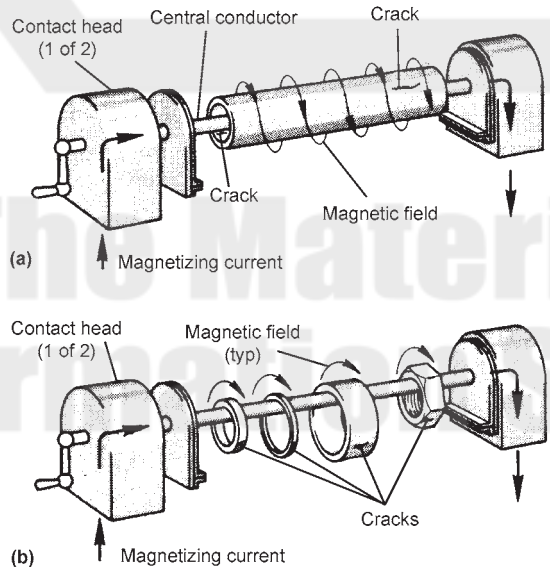


Fig. 8 Use of central conductors for circular magnetization of (a) long hollow cylindrical parts and (b) short hollow cylindrical and ring-like parts to detect discontinuities on inner and outer surfaces. Source: Ref 2

a convenient way to circularly magnetize a part without the need to make direct contact with the part itself. Central conductors are made of solid and tubular nonmagnetic and ferromagnetic materials that are good electrical conductors.

The basic rules regarding magnetic fields around a circular conductor carrying direct current are:

- The magnetic field outside a conductor of uniform cross-section is uniform along the length of the conductor
- The magnetic field is 90° to the path of the current through the conductor
- The flux density outside the conductor varies inversely with the radial distance from the center of the conductor

Direct Contact Method. For small parts having no openings through the interior, circular magnetic fields are produced by direct contact to the part. Parts are clamped between contact heads (head shot), generally on a bench unit (Fig. 9) that incorporates the source of current. A similar unit can be used to supply the magnetizing current to a central conductor (see Fig. 8).

Contact heads must be constructed so the surfaces of the part are not damaged–either physically by pressure, or structurally by heat from arcing and from high resistance at the points of contact. Such heat can be especially damaging to hardened surfaces such as bearing races.

For complete inspection of a complex part, it could be necessary to attach clamps at several points on the part or to wrap cables around the part to get fields in the proper directions at all points on the surface. This often

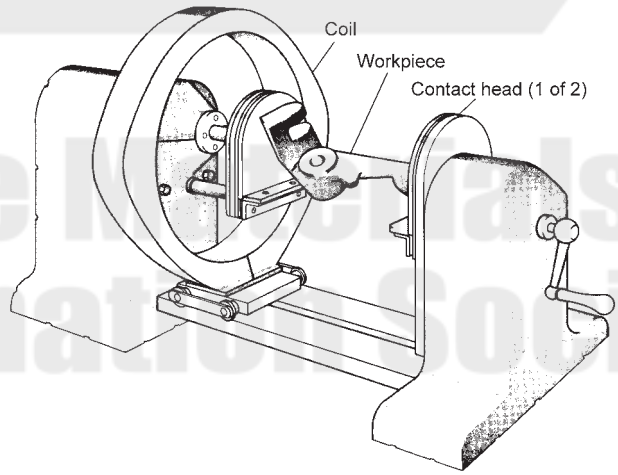


Fig. 9 Bench unit used to circularly magnetize workpieces clamped between contact heads (direct contact, head shot method). The coil on the unit can be used for longitudinal magnetization. Source: Ref 2

requires several magnetizations. The need for multiple magnetizations can be minimized by using the overall magnetization method, multidirectional magnetization, and induced current magnetization.

Prod Contacts. Magnetization often is done using prod contacts to inspect large and massive parts too bulky to be put into a unit having clamping contact heads. The method passes current directly through the part or through a local portion of it, as shown in Fig. 10. Such local contacts do not always produce true circular fields, but are very convenient and practical in many applications; e.g., prod contacts often are used in magnetic particle inspection of large castings and weldments.

Prod contacts have many advantages. Easy portability makes them convenient to use for field inspection of large tanks and welded structures. Sensitivity to defects lying wholly below the surface is greater with this method of magnetization than with any other, especially when half-wave current is used in conjunction with dry powder and the continuous method of magnetization.

Some limitations of using prod contacts are:

- Suitable magnetic fields exist only between and near the prod contact points. These points are seldom more than 300 mm (12 in.) apart, and usually much less. Therefore, it sometimes is necessary to relocate the prods so the entire surface of a part can be inspected
- Interference of the external field that exists between the prods sometimes makes observation of pertinent indications difficult. The strength of the current that can be used is limited by this effect
- Great care must be used to avoid burning the part under the contact points. Burning can be caused by dirty contacts, insufficient contact pressure, and excessive currents

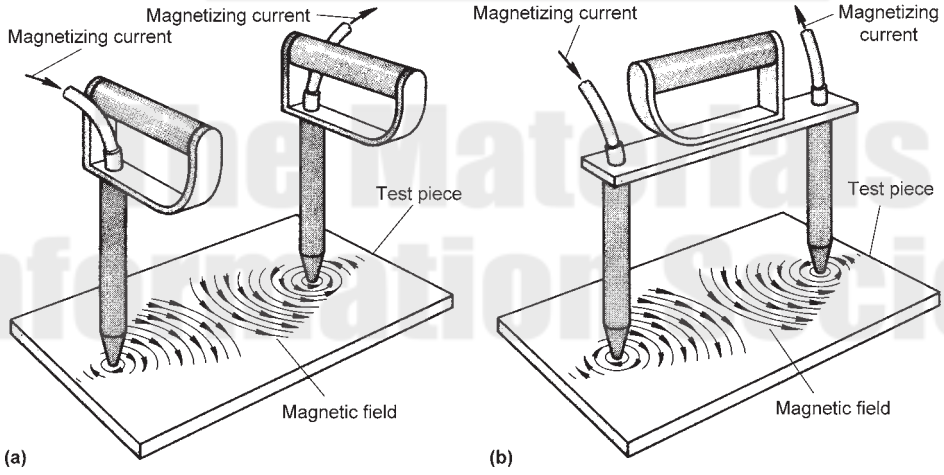


Fig. 10 (a) Single and (b) double prod contacts. Discontinuities are detected by a magnetic field generated between the prods. Source: Ref 2

Induced current provides a convenient method to generate circumferential magnetizing current ring shaped parts without making electrical contact. This is accomplished by properly orienting the ring within a magnetizing coil such that it links or encloses lines of magnetic flux (flux linkage), as shown in Fig. 11(a). As the level of magnetic flux changes (increases or decreases), a current flows around the ring in a direction opposing the change in flux level. The magnitude of this current depends on the total flux linkages, rate of flux linkage changes, and the electrical impedance associated with the current path within the ring. Increasing the flux linkages and the rate of change increases the magnitude of current induced in the ring. The circular field associated with this current takes the form of a toroidal magnetic field that encompasses all surface areas on the ring and that is conducive to revealing circumferential types of discontinuities. This is shown schematically in Fig. 11(b).

The choice of magnetizing current for the induced current method depends on magnetic properties of the part to be inspected. In instances where the residual method is applicable, such as for most bearing races and similar parts having high magnetic retentivity, direct current is used to

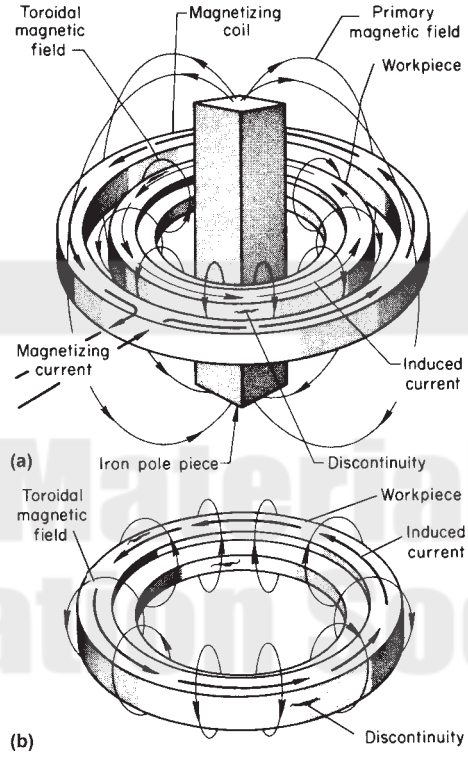


Fig. 11 Induced current method of magnetizing a ring shape part. (a) Ring being magnetized by induced current. Current direction corresponds to decreasing magnetizing current. (b) Resulting induced current and toroidal magnetic field in a ring. Source: Ref 2

magnetize. The rapid interruption of the current by quick break circuitry results in a rapid collapse of the magnetic flux and the generation of a high-amperage, circumferentially directed single pulse of current in the part. Thus, the part is residually magnetized with a toroidal field, and subsequent application of magnetic particles produces indications of circumferentially-oriented discontinuities.

A similar type of current of opposite polarity and lower amplitude is associated with the increasing flux due to the rapidly rising current, but in this case, only that current generated by the sudden breaking of the direct current serves a useful purpose.

Passing an alternating current through a conductor generates a fluctuating magnetic field as the level of magnetic flux rapidly changes from a maximum value in one direction to an equal value in the opposite direction. This is similar to the current that flows in a single shorted turn secondary of a transformer. The alternating induced current in conjunction with the continuous method renders the method applicable for processing magnetically soft, or less retentive, parts.

The induced current method, in addition to eliminating the possibility of damaging the part, is also capable of magnetizing in one operation parts that otherwise would require more than one head shot. Two examples of this type of part are illustrated in Fig. 12 and 13. These parts cannot be completely processed by one head shot to reveal circumferential defects, because regions at the contact points are not properly magnetized. Therefore, a two-step inspection process is required for full coverage, with the part rotated approximately 90° prior to the second step. Conversely, the induced current method provides full coverage in one processing step. The disk shaped part shown in Fig. 13 presents an additional problem when the contact method is used to reveal circumferential defects in the vicinity of the rim. Even when a two-step process is used, as with the ring in Fig. 12, the primary current path through the part might not develop a circular field of sufficient magnitude in the rim area. The induced current

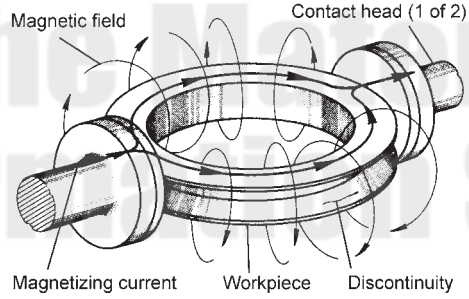


Fig. 12 Current and magnetic field distribution in a ring being magnetized with a head shot. Because regions at contact points are not magnetized, two operations are required for full coverage. With use of the induced current method, parts of this shape can be completely magnetized in one operation. Source: Ref 2

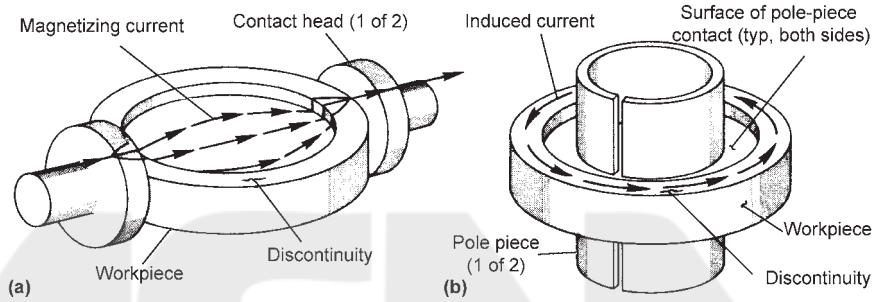


Fig. 13 Current paths in a rimmed disk shaped part magnetized by (a) head shot magnetization, and (b) induced current magnetization. Source: Ref 2

can be selectively concentrated in the rim area by proper pole piece selection to provide full coverage (rim area) in a single processing step. The pole pieces depicted in Fig. 13(b) are hollow and cylindrical, with one on each side of the disk. The pole pieces direct the magnetic flux through the disk such that the rim is the only portion constituting a totally enclosing current path.

Pole pieces used in conjunction with this method are preferably constructed of laminated ferromagnetic material. This minimizes the flow of eddy currents within the pole pieces, which detract from the induced (eddy) current developed within the part being processed. Pole pieces also can be made of rods, wire filled nonconductive tubes, and thick wall pipe, saw cut to break up the eddy current path.

Magnetic Particles and Suspending Liquids

Magnetic particles are classified according to the vehicle by which they are carried to the part: by air (dry particle method) and by a liquid (wet particle method). Magnetic particles consist of fine iron; black, brown, and red iron oxide (magnitite Fe_3O_4); brown iron oxide (γ-Fe_2O_3), ferrospinel ferrites ($Ni_xFe_2O_4$), and some nickel alloys. Important particle characteristics include magnetic properties, size, shape, density, mobility, and degree of visibility and contrast.

Magnetic Properties. Particles used for magnetic particle inspection should have high magnetic permeability so they can be readily magnetized by the low level leakage fields that occur around discontinuities and can be drawn by these fields to discontinuities to form readable indications. (The fields at very fine discontinuities can be extremely weak.) Particles also should have low coercive force and low retentivity.

Effect of Particle Size. Large, heavy particles are not likely to be arrested and held by weak fields when moving over a part surface, but fine particles are held by very weak fields. However, extremely fine particles can adhere to surface areas where there are no discontinuities (especially

if the surface is rough) and form confusing backgrounds. Coarse, dry particles fall too fast and are likely to bounce off the part surface without being attracted by the weak leakage fields at imperfections. Finer particles can adhere to fingerprints, rough surfaces, and soiled or damp areas, thereby obscuring indications.

Effect of Particle Shape. Particle shape can be spherical, needlelike, and rod like in form. Elongated needles tend to develop into little magnets with north-south poles, and, therefore, form into distinct, well defined patterns, which provide a more clear indication of the presence of a weak magnetic field. However, there is an optimum elongation aspect ratio for particles. The ability of dry particles to flow freely and to form uniformly dispersed clouds of powder that will spread evenly over a surface is a necessary characteristic for rapid and effective dry powder testing. The behavior of wet powder (suspensions) is less dependent on particle shape.

Visibility and contrast are enhanced by using particles with colors that make them easy to see against the color of the surface of the part being inspected. The natural color of the metallic powders used in the dry method is silver gray, but pigments are used to color them. The colors of particles for the wet method are limited to the black and red of the iron oxides commonly used as the base for wet particles.

For increased visibility, particles are coated with fluorescent pigment by the manufacturer. Inspection is conducted in total or partial darkness, using ultraviolet light to activate the fluorescent dyes. Inspected surfaces should be illuminated with a minimum of 1000 $\mu W/cm^2$ of black light, with a maximum of 2 ftc of general visible light at the inspection station. Fluorescent magnetic particles are available for use with both wet and dry methods. The fluorescent wet method is more common.

Dry particles are available in a variety of colors, some of them fluorescent. Color contrast powders should be viewed in ordinary light of a minimum of 100 ftc at the inspection station. Dry particles are most sensitive for use on very rough surfaces and for detecting flaws beneath the surface. They are ordinarily used with portable equipment. Reclaiming and reusing dry particles is not recommended.

Wet particles are best suited for detection of fine discontinuities such as fatigue cracks. Wet particles commonly are used in stationary equipment where the bath can remain in use until contaminated or until the properties of the particles are exhausted. They also are used in field operations with portable equipment, but the bath should be agitated constantly.

Oil Suspending Liquid. The oil used as a suspending liquid for magnetic particles should be an odorless, well-refined, light petroleum distillate of low viscosity and a high flash point. Oil viscosity should not exceed 0.03 cm^2/s when tested at 38 °C (100 °F), and must not exceed 0.05 cm^2/s when tested at the temperature prevailing at the point on the part being inspected. Above 0.05 cm^2/s, the movement of magnetic particles in the bath is sufficiently retarded to have a definite effect in reducing

buildup; therefore, reducing visibility of an indication of a small discontinuity. Parts should be precleaned to remove oil and grease because oil from the surface builds up in the bath and increases its viscosity.

Water Suspending Liquid. The use of water instead of oil for magnetic particle, wet method baths reduces costs and eliminates bath flammability. Water suspendible particle concentrates include the necessary wetting agents, dispersing agents, rust inhibitors, and antifoam agents.

Strength of the bath is a major factor in determining the quality of the indications obtained. The proportion of magnetic particles in the bath must be maintained at a uniform level. The strength of indications varies with varying concentration, which could cause misinterpretation of indications. Fine indications can be missed entirely with a weak bath. High concentrations produce a confusing background and excessive adherence of particles at external poles, which interferes with distinct indications of extremely fine discontinuities.

The best method to ensure optimum bath concentration for any given combination of equipment, bath application, and type of part and discontinuities involved, is to test the bath using parts with known discontinuities. Bath strength can be adjusted until satisfactory indications are obtained. This bath concentration can then be adopted as standard for similar conditions.

Bath concentration can be measured reasonably accurately using the settling test. In the test, 100 mL (0.03 gal) of well agitated bath is placed in a pear shape centrifuge tube. The volume of solid material that settles out after a predetermined interval (usually 30 minutes) is measured on the graduated cylindrical part of the tube. Dirt in the bath also will settle and usually shows as a separate layer on top of the oxide. The layer of dirt usually is easily distinguishable because it is different in color from the magnetic particles.

Ultraviolet Light

A mercury vapor lamp is a convenient source of ultraviolet light, emitting a light spectrum that has several intensity peaks within a wide band of wavelengths. When used for a specific purpose, emitted light is passed through a suitable filter so only a relatively narrow band of ultraviolet wavelength is available. For example, a band in the long wave ultraviolet spectrum is used for fluorescent liquid penetrant or magnetic particle inspection.

Fluorescence is the quality of an element or combination of elements to absorb the energy of light at one frequency and emit light of a different frequency. Fluorescent materials used in liquid penetrant and magnetic particle inspection are combinations of elements selected to absorb light in the peak energy band of the mercury vapor lamp fitted with a Kopp glass filter. This peak occurs at about 365 nm (14.4 μin.). The ability of fluorescent materials to emit light in the greenish-yellow wavelengths of

the visible spectrum depends on the intensity of ultraviolet light at the workpiece surface.

Detectable Discontinuities

The usefulness of magnetic particle inspection in the search for discontinuities or imperfections depends on exactly what types of discontinuities the method is capable of finding. Of importance are the size, shape, orientation, and location of the discontinuity, with respect to its ability to produce leakage fields.

Surface Discontinuities. The largest and most important type of discontinuities consists of those that are exposed to the surface. Surface cracks or discontinuities are effectively located using magnetic particles. Surface cracks, such as those shown in Fig. 14, are also more detrimental to the service life of a component than are subsurface discontinuities, and, therefore, they are more frequently the object of inspection.

Magnetic particle inspection is capable of locating seams, laps, quenching and grinding cracks, and surface ruptures in castings, forgings, and weldments. The method also can detect surface fatigue cracks developed during service. Magnetizing and particle application methods can be critical in certain instances, but in most applications the requirements are relatively easily met, because leakage fields usually are strong and highly localized.

To successfully detect a discontinuity, there must be a field of sufficient strength in a generally favorable direction to produce strong leakage fields. For maximum detectability, the field generated in the part should be at right angles to the length of a suspected discontinuity (see Fig. 6). This is especially true if the discontinuity is small and fine.

Subsurface discontinuities comprise those voids or nonmetallic inclusions that lie just beneath the surface. Nonmetallic inclusions are present in all steel products to some degree. They occur as scattered individual

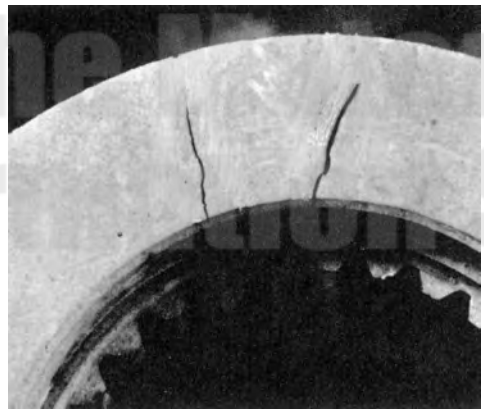


Fig. 14 Magnetic particle indications of cracks in a large cast splined coupling. Source: Ref 2

inclusions, or they may be aligned in long stringers. These discontinuities usually are very small and cannot be detected unless they lie very close to the surface because they produce highly localized but rather weak fields.

Nonrelevant Indications

Nonrelevant indications are true patterns caused by leakage fields that do not result from the presence of flaws. The term *false indications* is sometimes used to describe this type of indication, because the indication falsely implies the presence of a flaw, even though the particle buildup actually results from a leakage field. There are several possible causes of nonrelevant indications, which require evaluation but should not be interpreted as flaws.

Demagnetization after Inspection

All ferromagnetic materials retain a residual magnetic field to some degree after being magnetized. This field is negligible in magnetically soft metals, but in harder metals, it can be comparable to the intense fields associated with the special alloys used for permanent magnets.

It is not always necessary to demagnetize parts, but it is essential in many cases, even though it is costly and time consuming. The degree of difficulty in demagnetization depends on the type of metal. Metals having high coercive force are the most difficult to demagnetize. High retentivity is not necessarily related directly to high coercive force, so the strength of the retained magnetic field is not always an accurate indicator of the ease of demagnetizing.

There are several reasons to demagnetize a part after magnetic particle inspection, or after any other magnetization. Demagnetize if:

- The part is used in an area where a residual magnetic field interferes with the operation of instruments sensitive to magnetic fields, or where it can affect the accuracy of instrumentation incorporated in an assembly that contains the magnetized part
- Chips might adhere to the surface during subsequent machining and adversely affect surface finish, dimensions, and tool life
- Chips might adhere to the surface during cleaning operations and interfere with subsequent operations such as painting and plating
- Abrasive particles might be attracted to magnetized parts, such as bearing surfaces, bearing raceways, and gear teeth, resulting in abrasion and galling, and obstruction of oil holes and grooves
- Strong residual magnetic fields can deflect the arc away from the point at which it should be applied during some arc welding operations
- A residual magnetic field in a part can interfere with remagnetization of the part at a field intensity too low to overcome the remanent field in the part

Demagnetization might not be necessary if:

- Parts are made of magnetically soft steel having low retentivity; such parts usually will become demagnetized as soon as they are removed from the magnetizing source
- The parts are subsequently heated above their Curie point and consequently lose their magnetic properties
- The magnetic field is such that it will not affect the function of the part in service
- The part is to be remagnetized for further magnetic particle inspection or for some secondary operation in which a magnetic plate or chuck may be used to hold the part.

The last reasons for demagnetizing and not demagnetizing seem to be contradictory. The establishment of a longitudinal field after circular magnetization negates the circular field because two fields in different directions cannot exist in the same part at the same time. If the magnetizing force is not of sufficient strength to establish the longitudinal field it should be increased, or other steps taken to ensure that the longitudinal field actually has been established. The same is true in changing from longitudinal to circular magnetization. If the two fields (longitudinal and circular) are applied simultaneously, a field is established that is a vector combination of the two in both strength and direction. However, if the fields are impressed successively, the last field applied, if strong enough to establish itself in the part, destroys the remanent field from the previous magnetization.

Eddy Current Inspection

Eddy current inspection is based on the principles of electromagnetic induction and is used to identify or differentiate a wide variety of physical, structural, and metallurgical conditions in electrically conductive ferromagnetic and nonferromagnetic metals and metal parts. Eddy current inspection is used:

- To measure and identify conditions and properties related to electrical conductivity, magnetic permeability, and physical dimensions (primary factors affecting eddy current response)
- To detect seams, laps, cracks, voids, and inclusions
- To sort dissimilar metals and detect differences in their composition, microstructure, and other properties (such as grain size, heat treatment, and hardness)
- To measure the thickness of a nonconductive coating on a conductive metal, or the thickness of a nonmagnetic metal coating on a magnetic metal

Because eddy current inspection is an electromagnetic induction technique, it does not require direct electrical contact with the part being inspected. The eddy current method is adaptable to high speed inspection, and because it is nondestructive, it can be used to inspect an entire production output if desired. The method is based on indirect measurement, and the correlation between instrument readings and the structural characteristics and serviceability of parts being inspected must be carefully and repeatedly established.

Eddy current inspection is extremely versatile, which is both an advantage and a disadvantage. The advantage is that the method can be applied to many inspection problems provided that the physical requirements of the material are compatible with the inspection method. However, in many applications, the sensitivity of the method to many inherent material properties and characteristics can be a disadvantage. Some variables in a material that are not important in terms of material or part serviceability can cause instrument signals that mask critical variables or are mistakenly interpreted to be caused by critical variables.

Eddy Current Versus Magnetic Inspection Methods. In eddy current inspection, eddy currents create their own electromagnetic field, which is sensed either through the effects of the field on the primary exciting coil or by means of an independent sensor. In nonferromagnetic materials, the secondary electromagnetic field is derived exclusively from eddy currents. However, with ferromagnetic materials, additional magnetic effects occur that usually are of sufficient magnitude to overshadow the basic eddy current effects from electrical conductivity only. These magnetic effects result from the magnetic permeability of the material being inspected, and can be virtually eliminated by magnetizing the material to saturation in a static (direct current) magnetic field. When the permeability effect is not eliminated, the inspection method is more correctly categorized as electromagnetic or magnetoinductive inspection.

Principles of Operation

Functions of a Basic System. The part to be inspected is placed within or adjacent to an electrical coil in which an alternating current is flowing. As shown in Fig. 15, the alternating current, called the exciting current, causes eddy currents to flow in the part as a result of electromagnetic induction. These currents flow within closed loops in the part, and their magnitude and timing (or phase) depend on (a) the original or primary field established by the exciting currents, (b) the electrical properties of the part, and (c) the electromagnetic fields established by currents flowing within the part.

The electromagnetic field in the region in the part and surrounding the part depends on both the exciting current from the coil and the eddy currents flowing in the part. The flow of eddy currents depends on the electri-

cal characteristics of the part, the presence or absence of flaws and other part discontinuities, and the total electromagnetic field within the part.

The change in flow of eddy currents caused by the presence of a crack in a pipe is shown in Fig. 16. The pipe travels along the length of the inspection coil, as shown. In section A-A in Fig. 16, no crack is present and the eddy current flow is symmetrical. In section B-B, where a crack is

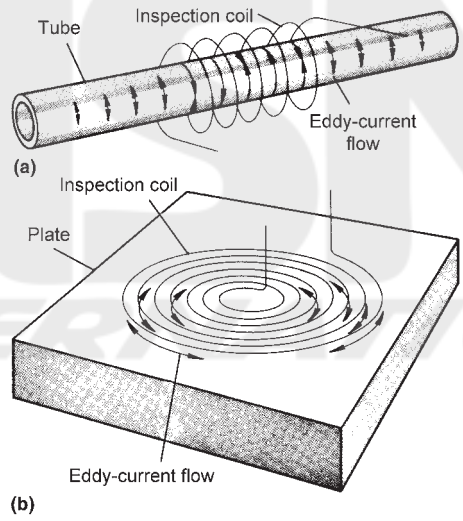


Fig. 15 Two common types of inspection coils and the patterns of eddy current flow generated by the exciting current in the coils. (a) Solenoid type coil is applied to cylindrical or tubular parts. (b) Pancake type coil applied to a flat surface. Source: Ref 3

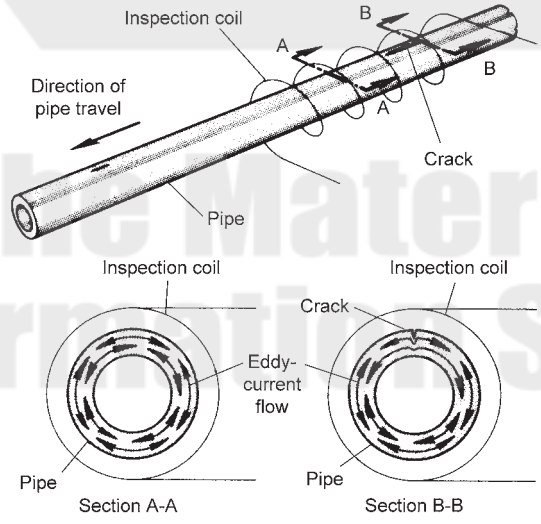


Fig. 16 Effect of a crack on the pattern of eddy current flow in a pipe. Source: Ref 3

present, the eddy current flow is impeded and changed in direction, causing significant changes in the associated electromagnetic field. The condition of the part can be monitored by observing the effect of the resulting field on the electrical characteristics of the exciting coil, such as its electrical impedance, induced voltage, and induced currents. Alternatively, the effect of the electromagnetic field can be monitored by observing the induced voltage in one or more other coils placed within the field near the part being monitored.

Each and all of these changes can have an effect on the exciting coil and other coil or coils used to sense the electromagnetic field adjacent to a part. The effects most often used to monitor the condition of the part being inspected are the electrical impedance of the coil and the induced voltage of either the exciting coil or other adjacent coil or coils.

Eddy current systems vary in complexity depending on individual inspection requirements. However, most systems must provide for the following functions:

- Excitation of the inspection coil with one or more frequencies
- Modulation of the inspection coil output signal by the part being inspected
- Processing of the inspection coil signal prior to amplification
- Amplification of the inspection coil signals
- Detection or demodulation of the inspection coil signal, usually accompanied by some analysis or discrimination of signals, which can be performed by a computer
- Display of signals on an instrument such as a meter, an oscilloscope, an oscillograph, and a strip chart recorder; or recording of signals on paper punch tape and magnetic tape
- Handling of the part being inspected and support of inspection coil assembly

Elements of a typical inspection system are shown schematically in Fig. 17. The particular elements in Fig. 17 are for a system developed to inspect bar or tubing. The generator supplies excitation current to the inspection coil and a synchronizing signal to the phase shifter, which provides switching signals for the detector. The loading of the inspection coil by the part being inspected modulates the electromagnetic field of the coil. This causes changes in the amplitude and phase of the inspection coil voltage output.

The output of the inspection coil is fed to the amplifier and detected or demodulated by the detector. The demodulated output signal, after some further filtering and analyzing, is then displayed on an oscilloscope or a chart recorder. The displayed signals, having been detected or demodulated, vary at a much slower rate, depending on (a) the rate of changing the inspection probe from one part being inspected to another; (b) the

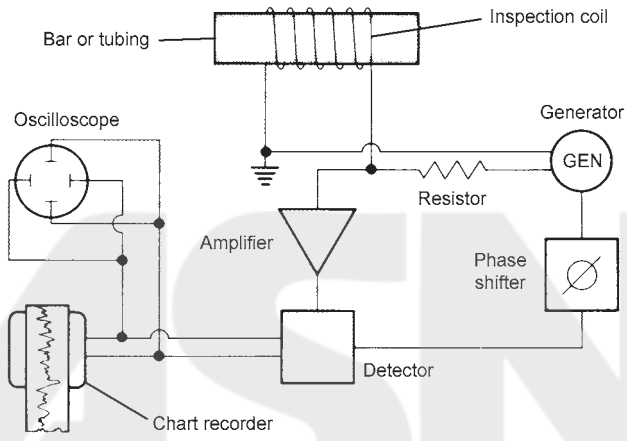


Fig. 17 Principal elements of a typical system for eddy current inspection of bar or tubing. See description in text. Source: Ref 3

speed at which the part is fed through an inspection coil; or, (c) the speed at which the inspection coil is caused to scan past the part being inspected.

Operating Variables

The principal operating variables encountered in eddy current inspection include coil impedance; electrical conductivity; magnetic permeability; lift-off and fill factors; edge effect; and skin effect.

Coil Impedance. When direct current flows in a coil, the magnetic field reaches a constant level and the electrical resistance of the wire is the only limitation to the flow of current. However, when alternating current flows in a coil, two limitations are imposed: the alternating current resistance of the wire and a quantity known as inductive reactance (X_L).

Impedance usually is plotted on an impedance plane diagram. In such a diagram, resistance is plotted along one axis and inductive reactance (or inductance) along the other axis. Because each specific condition in the material being inspected can result in specific coil impedance, each condition corresponds to a particular point on the impedance plane diagram. For example, if a coil is placed sequentially on a series of thick pieces of metal, each having a different resistivity, each piece causes different coil impedance and corresponds to a different point on a locus in the impedance plane. The curve generated might resemble that shown in Fig. 18, which is based on International Annealed Copper Standard (IACS) conductivity ratings. Other curves are generated for other material variables, such as section thickness and types of surface flaws. By use of more than one test frequency, the impedance planes can be manipulated to accept a desirable variable (in flaws) and reduce the effects of undesirable variables–that is, lift-off and/or dimensional effects.

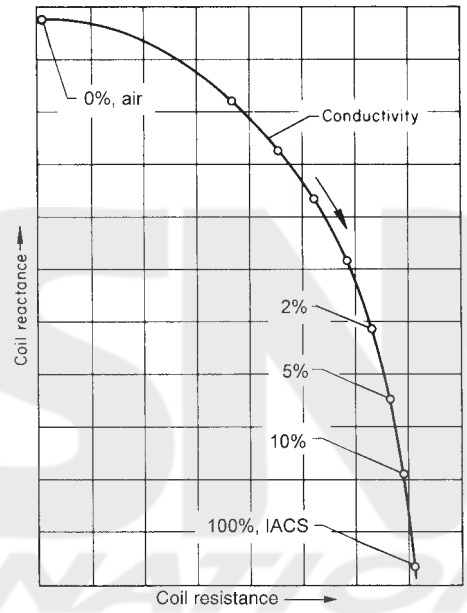


Fig. 18 Typical impedance plane diagram derived by placing an inspection coil sequentially on a series of thick pieces of metal, each with a different International Annealed Copper Standard (IACS) electrical resistance or conductivity rating. The inspection frequency was 100 kHz. Source: Ref 3

Electrical Conductivity. All materials have a characteristic resistance to the flow of electricity. Those with the highest resistivity are classified as insulators; those having intermediate resistivity are classified as semiconductors; and, those having low resistivity are classified as conductors. Conductors, which include most metals, are of greatest interest in eddy current inspection. The relative conductivities of common metals and alloys vary over a wide range.

The capacity to conduct current is measured in terms of either conductivity or resistivity. In eddy current inspection, measurement often is based on IACS. In this system, the conductivity of annealed, unalloyed copper is arbitrarily rated at 100%, and the conductivities of other metals and alloys are expressed as percentages of this standard. Thus, the conductivity of unalloyed aluminum is rated 61% IACS, or 61% that of unalloyed copper. The resistivities and IACS conductivity ratings of several common metals and alloys are given in Table 3.

Magnetic Permeability. Ferromagnetic metals and alloys, including iron, nickel, cobalt, and some of their alloys, concentrate the flux of a magnetic field. They are strongly attracted to a magnet and an electromagnet, have exceedingly high and variable susceptibilities, and have very high and variable permeabilities.

Magnetic permeability is not a constant for a given material, but depends on the strength of the magnetic field acting on it. For example, con-

sider a sample of steel that has been completely demagnetized and then placed in a solenoid coil. As current in the coil is increased, the magnetic field associated with the current increases. However, the magnetic flux within the steel increases rapidly at first and then levels off so that an additionally large increase in the strength of the magnetic field results in only a small increase in flux within the steel. The steel sample achieves a condition known as magnetic saturation.

The curve showing the relation between magnetic field intensity and the magnetic flux within the steel is known as a magnetization curve. Magnetization curves for annealed commercially pure iron and nickel are shown in Fig. 19. The magnetic permeability of a material is the ratio between the strength of the magnetic field and the amount of magnetic flux within the material. As shown in Fig. 19, at saturation (where there is no appre-

Table 3 Electrical resistivity and conductivity of several common metals and alloys

Metal or alloy	Resistivity, μΩ · mm	Conductivity, %IACS
Silver	16.3	105
Copper, annealed	17.2	100
Gold	24.4	70
Aluminum	28.2	61
Aluminum alloys		
6061-T6	41	42
7075-T6	53	32
2024-T4	52	30
Magnesium	46	37
70-30 brass	62	28
Phosphor bronzes	160	11
Monel	482	3.6
Zirconium	500	3.4
Zircaloy-2	720	2.4
Titanium	548	3.1
Ti-6Al-4V alloy	1720	1.0
Type 304 stainless steel	700	2.5
Inconel 600	980	1.7
Hastelloy X	1150	1.5
Waspaloy	1230	1.4

Source: Ref 3

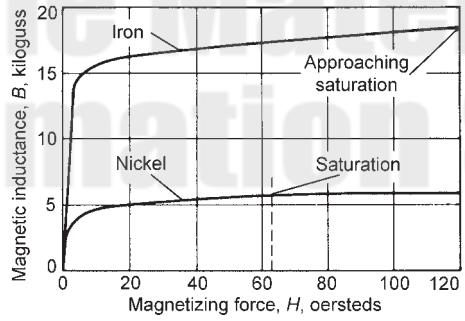


Fig. 19 Magnetization curves for annealed commercially pure iron and nickel. Source: Ref 3

ciable change in induced flux in the material for a change in field strength) the permeability is nearly constant for small changes in field strength. The magnetic permeability of the material being inspected strongly influences the eddy current response. Consequently, the techniques and conditions used for inspecting magnetic materials differ from those used to inspect nonmagnetic materials.

Lift-Off Factor. When a probe inspection coil, attached to a suitable inspection instrument, is energized in air, it produces an indication even if there is no conductive material in the vicinity of the coil. The initial indication starts to change as the coil is moved closer to a conductor. Because the field of the coil is strongest close to the conductor, the indicated change on the instrument continues to increase until the coil is directly on the conductor. These changes in indication with changes in spacing between the coil and the conductor, or part being inspected, are called *lift off*. The lift-off effect is so pronounced that small variations in spacing can mask many indications resulting from the condition or conditions of primary interest. Consequently, it usually is necessary to maintain a constant relationship between the size and shape of the coil and the size and shape of the part being inspected.

The change of coil impedance with lift-off can be derived from the impedance plane diagram shown in Fig. 20. When the coil is suspended in air away from the conductor, impedance is at a point at the upper end of the curve at far left in Fig. 20. As the coil approaches the conductor, the impedance moves in the direction indicated by the dashed lines until the coil is in contact with the conductor. When contact occurs, the impedance is at a point corresponding to the impedance of the part being inspected, which in this instance represents its conductivity. The fact that the lift-off curves approach the conductivity curve at an angle can be used in some instruments to separate lift-off signals from those resulting from variations in conductivity or some other parameter of interest.

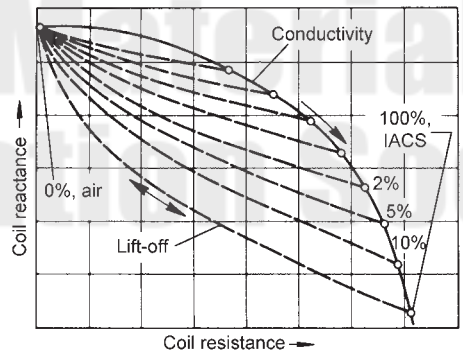


Fig. 20 Impedance plane diagram showing curves for electrical conductivity and lift off. Inspection frequency was 100 kHz. Source: Ref 3

Although lift off can be troublesome in many applications, it can be also useful. For example, using the lift-off effect, eddy current instruments are excellent for measuring the thickness of nonconductive coatings, such as paint and anodized coatings, on metals.

Fill Factor. In an encircling coil, a condition comparable to lift-off is known as *fill factor*. It is a measure of how well the part being inspected fills the coil. As with lift off, changes in fill factor resulting from factors such as variations in outside diameter must be controlled because small changes can produce large indications. The lift-off curves shown in Fig. 20 are very similar to those for changes in fill factor. For a given lift-off or fill factor, the conductivity curve shifts to a new position, as indicated in Fig. 20. Fill factor can sometimes be used as a rapid method to check variations in outside diameter measurements in rods and bars.

Edge Effect. When an inspection coil approaches the end or edge of a part, eddy currents are distorted because they are unable to flow beyond the edge of a part. The eddy current distortion of eddy results in an indication known as *edge effect*. Because the magnitude of the effect is very large, it limits inspection near edges. Unlike lift-off, little can be done to eliminate edge effect. A reduction in coil size somewhat reduces the effect, but there are practical limits that dictate the sizes of coils for given applications. In general, it is not advisable to inspect any closer than 3.2 mm (⅛ in.) from the edge of a part.

One alternative for inspection near an edge with minimal edge effect is to scan in a line parallel to the edge. Inspection can be carried out by maintaining a constant probe to edge relationship, but each new scan line position requires adjustment of the instrument. Fixturing of the probe is recommended.

Skin Effect. Eddy currents are not uniformly distributed throughout a part being inspected; rather, they are densest at the surface immediately beneath the coil and become progressively less dense with increasing distance below the surface. The concentration of eddy currents at the surface of a part is known as *skin effect*. At some distance below the surface of a thick part, there are essentially no currents flowing. The depth of eddy current penetration should be considered for thickness measurements and for detection of subsurface flaws.

Figure 21 shows how the eddy current varies as a function of depth below the surface. The depth at which the density of the eddy current is reduced to, about 37% of the density, at the surface is defined as the *standard depth of penetration*. This depth depends on the electrical conductivity and magnetic permeability of the material and on the frequency of the magnetizing current. Depth of penetration decreases with increases in conductivity, permeability, and inspection frequency. The standard depth of penetration can be calculated from the equation:

$$S = 1980\sqrt{\rho/\mu f}$$

where *S* is standard depth of penetration, in inches; ρ is resistivity, in ohm-centimeters; μ is magnetic permeability (1 for nonmagnetic materials); and *f* is inspection frequency, in hertz (Hz). The standard depth of penetration is shown in Fig. 22, as a function of inspection frequency, for several metals of various electrical conductivities.

Inspection Frequencies

The inspection frequencies used in eddy current inspection range from about 60 Hz to 6 MHz. Most inspection of nonmagnetic materials is per-

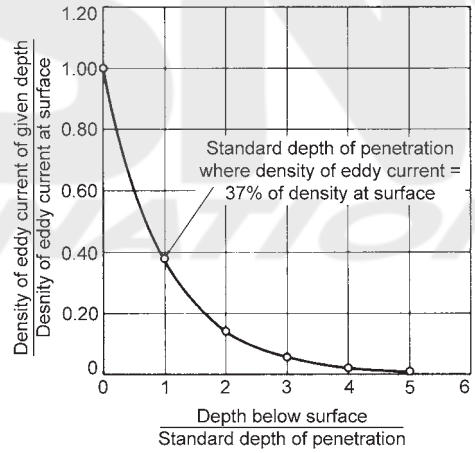


Fig. 21 Variation in density of eddy current as a function of depth below the surface of a conductor, known as skin effect. Source: Ref 3

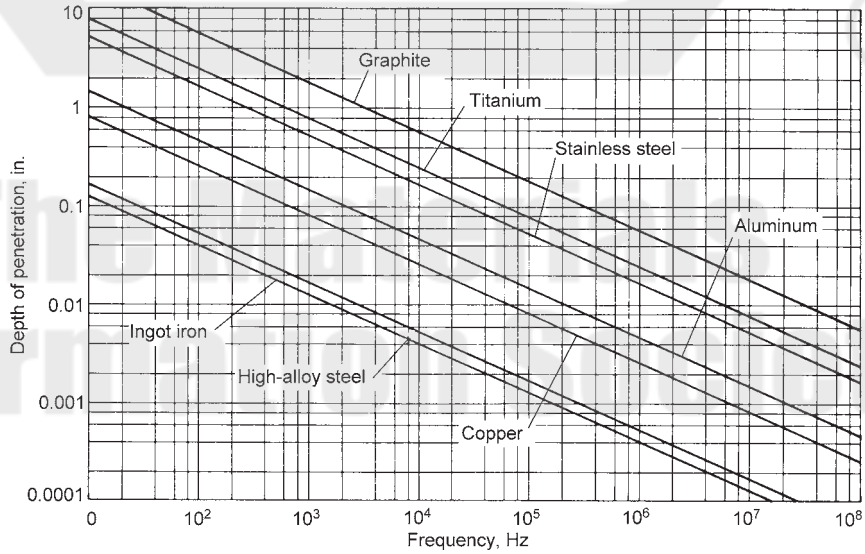


Fig. 22 Standard depths of penetration as a function of frequencies used in eddy current inspection for several metals of various electrical conductivities. Source: Ref 3

formed at a few kilohertz (kHz). In general, lower frequencies are used to inspect magnetic materials. However, the actual frequency used in any specific eddy current inspection depends on the thickness of the material being inspected, the required depth of penetration, the degree of sensitivity or resolution required, and the purpose of the inspection.

Selection of inspection frequency is normally a compromise. For example, penetration should be sufficient to reach subsurface flaws that must be detected, and to determine material condition such as case hardness. Although penetration is greater at lower frequencies, it does not follow that the lowest possible frequency should be used. Unfortunately, as the frequency is lowered, the sensitivity to flaws decreases somewhat and the speed of inspection could be reduced.

Typically, the highest possible inspection frequency that still is compatible with the penetration depth required is selected. The choice is relatively simple when only surface flaws must be detected, in which case frequencies up to several megahertz (MHz) can be used. However, when flaws at some considerable depth below the surface must be detected, or when flaw depth and size must be determined, low frequencies must be used at the expense of sensitivity.

In inspection of ferromagnetic materials, relatively low frequencies are typically used because of the low penetration in these materials. Higher frequencies can be used when it is necessary to inspect for surface conditions only. However, even the higher frequencies used in these applications are still considerably lower than those used to inspect nonmagnetic materials for similar conditions.

Inspection Coils

The inspection coil is an essential part of every eddy current inspection system. The shape of the inspection coil depends to a considerable extent on the purpose of the inspection and on the shape of the part being inspected. In inspection for flaws, such as cracks and seams, it is essential that the flow of the eddy currents be as nearly perpendicular to the flaws as possible to obtain a maximum response from the flaws. If the eddy current flow is parallel to flaws, there is little or no distortion of the currents; and, therefore, very little reaction on the inspection coils.

Probe and Encircling Coils. Of the almost infinite variety of coils used in eddy current inspection, probe coils and encircling coils are the most common. A probe type coil typically is used to inspect a flat surface for cracks at an angle to the surface because this type of coil induces currents that flow parallel to the surface; and, therefore, across a crack as shown in Fig. 23(a). Conversely, a probe type coil is not suitable to detect a laminar type of flaw. For such a discontinuity, a U-shape, or horseshoe shaped coil such as the coil shown in Fig. 23(b) is satisfactory.

To inspect tubing and bar, an encircling coil (Fig. 23c) is generally used because of complementary configuration and because of the testing speeds

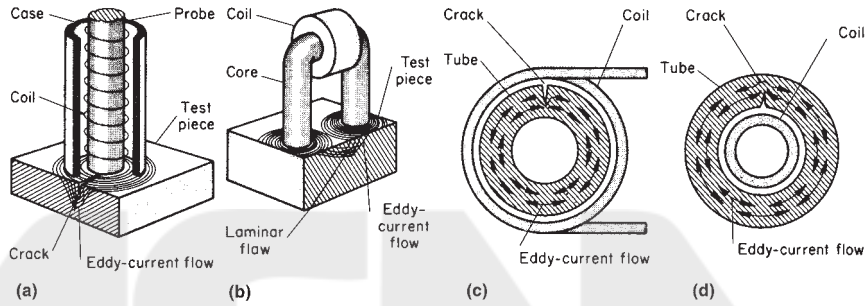


Fig. 23 Types and applications of coils used in eddy current inspection. (a) Probe type coil applied to a flat plate for crack detection. (b) Horseshoe shape, or U-shape, coil applied to a flat plate for laminar flaw detection. (c) Encircling coil applied to a tube. (d) Internal, or bobbin type, coil applied to a tube. Source: Ref 3

that can be achieved. However, an encircling coil is sensitive only to discontinuities that are parallel to the axis of the tube and bar. The coil is satisfactory for this particular application because most discontinuities in tubing and bar are parallel to the major axis as a result of the manufacturing process. If it is necessary to locate discontinuities that are not parallel to the axis, a probe coil must be used, and either the coil or the part must be rotated during scanning.

To detect discontinuities on the inside surface of a tube, an internal, or bobbin type, coil (Fig. 23d) can be used. An alternative is to use an encircling coil with a depth of penetration sufficient to detect flaws on the inside surface. The bobbin type coil, similar to the encircling coil, is sensitive to discontinuities that are parallel to the axis of the tube or bar.

Multiple Coils. In many eddy current inspection setups, two coils are used. The two coils are typically connected in a series opposing arrangement so there is no output from the pair when their impedances are the same. Pairs of coils can be used in either an absolute or a differential arrangement (Fig. 24). In the absolute arrangement (Fig. 24a), a sample of acceptable material is placed in one coil, and the other coil is used for inspection. In this manner, the coils compare an unknown against a standard; the differences between the two (if any) are indicated by a suitable instrument. Arrangements of this type are commonly used in sort applications. Fixtures are used to maintain a constant geometrical relationship between coil and part.

An absolute coil arrangement is not a good method in many applications. For example to inspect tubing, an absolute arrangement indicates dimensional variations in both outside diameter and wall thickness even though such variations can be well within allowable limits. To avoid this problem, a differential coil arrangement such as that shown in Fig. 24(b) can be used. Here, the two coils compare one section of the tube with an adjacent section. When the two sections are the same, there is no output from the pair of coils and no indication on the eddy current instrument.

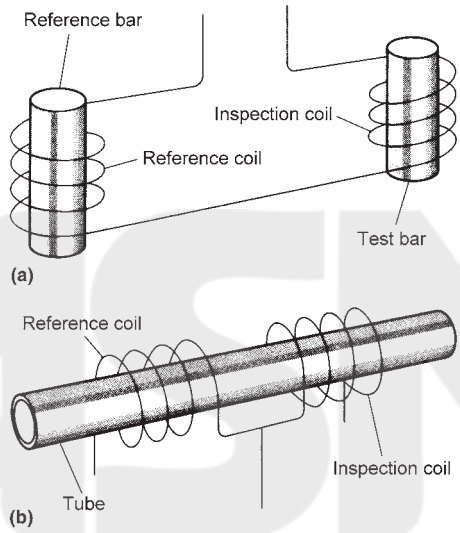


Fig. 24 Multiple coils used in eddy current inspection. (a) Absolute coil arrangement. (b) Differential coil arrangement. Source: Ref 3

Gradual dimensional variations within the tube or gross variations between individual tubes are not indicated, whereas discontinuities, which normally occur abruptly, are very apparent. In this way, it is possible to have an inspection system that is sensitive to flaws and relatively insensitive to changes that normally are not of interest.

Sizes and Shapes. Inspection coils are made in a variety of sizes and shapes. The selection of a coil for a particular application depends on the type of discontinuity. For example, when an encircling coil is used to inspect tubing and bar for short discontinuities, the best resolution is obtained with a short coil. On the other hand, a short coil has the disadvantage of being sensitive to the position of the part in the coil. Longer coils are not as sensitive to position of the part but are not as effective in detecting very small discontinuities. Small diameter probe coils have greater resolution than larger ones but are more difficult to manipulate and are more sensitive to lift-off variations.

Eddy Current Instruments

A simple eddy current instrument, in which the voltage across an inspection coil is monitored, is shown in Fig. 25(a). This circuit is adequate to measure large lift-off variations, if accuracy is not of great importance. A circuit designed for greater accuracy is shown in Fig. 25(b). This instrument consists of a signal source, an impedance bridge with dropping resistors, an inspection coil in one leg, and a balancing impedance in the other leg. The differences in voltage between the two legs of the bridge are measured by an alternating current voltmeter. Alternatively, the balancing im-

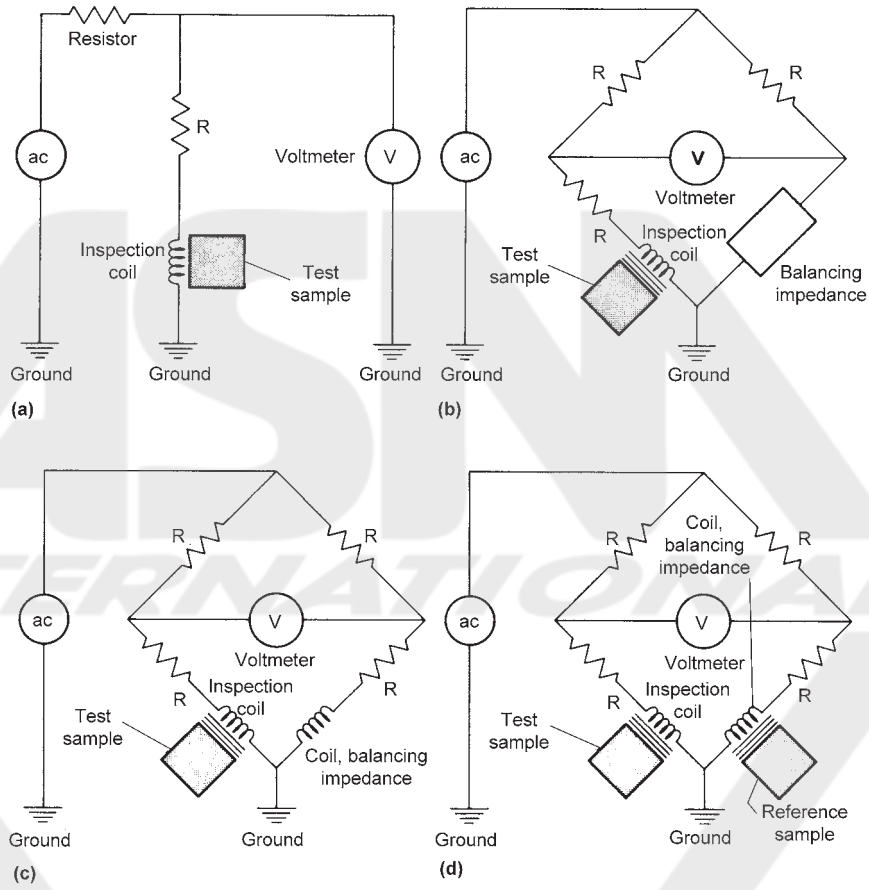


Fig. 25 Four types of eddy current instruments. (a) A simple arrangement, in which voltage across the coil is monitored. (b) Typical impedance bridge. (c) Impedance bridge with dual coils. (d) Impedance bridge with dual coils and a reference sample in the second coil. Source: Ref 3

pedance in the leg opposite the inspection coil can be a coil identical to the inspection coil, as shown in Fig. 25(c), or it can have a reference sample in the coil, as shown in Fig. 25(d). In the latter case, if all the other components in the bridge are identical, a signal occurs only when the inspection coil impedance deviates from that of the reference sample.

There are other methods to achieve bridge balance, such as varying the values of resistance of the resistor in the upper leg of the bridge and one in series with the balancing impedance. The most accurate bridges can measure absolute impedance to within 0.01%. However, in eddy current inspection, it is not how an impedance bridge is balanced that is important, but rather how it is unbalanced by the effects of a flaw.

Another type of bridge system is an induction bridge, in which the power signal is transformer coupled into an inspection coil and a reference coil. In addition, the entire inductance balance system is placed in the probe, as shown in Fig. 26. The probe consists of a large transmitter or driver coil and two small detector or pickup coils wound in opposite directions as mirror images of each other. An alternating current is supplied to the large transmitter coil to generate a magnetic field. If the transmitter coil is not in the vicinity of a conductor, the two detector coils detect the same field, and the net signal is zero because they are wound in opposition to each other. However, if one end of the probe is placed near a metal surface, the field is different at the two ends of the probe, and a net voltage appears across the two coils. The resultant field is the sum of a transmitted signal that is present all the time, and a reflected signal due to the presence of a conductor (the metal surface). This coil arrangement can be used both as a probe and as an encircling coil.

Readout Instrumentation. An important part of an eddy current inspection system is the instrument used for a readout. The readout device can be an integral part of the system, an interchangeable plug-in module,

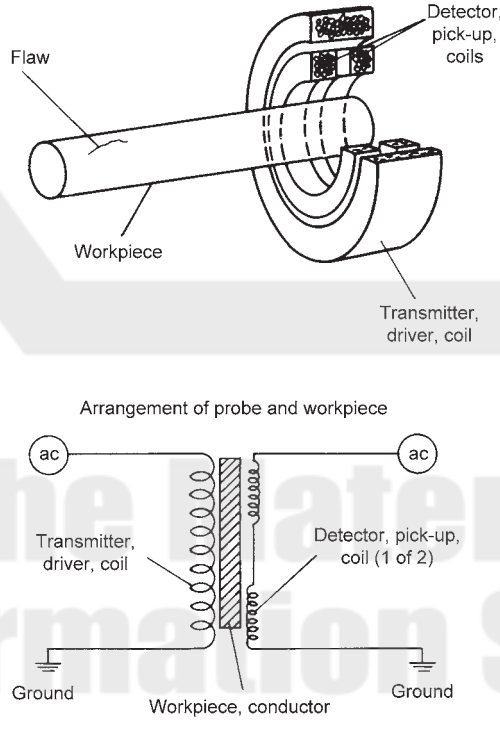


Fig. 26 Induction bridge probe in place at the surface of a workpiece. Schematic shows how power signal is transformer coupled from a transmitter coil into two detector coils—an inspection coil (at bottom) and a reference coil (at top). Source: Ref 3

and a solitary unit connected by cable. The readout instrument should be of adequate speed, accuracy, and range to meet the inspection requirements of the system. Frequently, several readout devices are used in a single inspection system. More common types of readout, in order of increasing cost and complexity, are:

- Alarm lights alert the operator that a test parameter limit has been exceeded
- Sound alarms serve the same purpose as alarm lights but free the attention of the operator to allow manipulating the probe in manual scanning.
- Kick-out relays activate a mechanism that automatically rejects and marks a part when a test parameter is exceeded
- Analog meters give a continuous reading over an extended range. They are fairly rapid (with a frequency of about 1 Hz), and the scales can be calibrated to read parameters directly. The accuracy of these devices is limited to about 1% of full scale. They can be used to set the limits on alarm lights, sound alarms, and kick-out relays
- Digital meters are easier to read and can have greater ranges than analog meters. Numerical values are easily read without extrapolation, but fast trends of changing readings are more difficult to interpret. Although many digital meters have binary coded decimal (bcd) output, they are relatively slow
- X-Y plotters can be used to display impedance plane plots of the eddy current response. They are very helpful in the design and set up of eddy current, bridge unbalance inspections and in discriminating against undesirable variables. They also are useful to sort out inspection results. They are fairly accurate and provide a permanent copy
- X-Y storage oscilloscopes are very similar to X-Y plotters but can acquire signals at high speed. However, the signals have to be processed manually, and the screen can quickly become cluttered with signals. In some instruments, high-speed X-Y gates can be displayed and set on the screen
- Strip chart recorders furnish a fairly accurate (about 1% of full scale) recording at reasonably high speed (about 200 Hz). However, once on the chart, the data must be read by an operator. Several channels can be recorded simultaneously, and the record is permanent
- Magnetic tape recorders are fairly accurate and capable of recording at very high speed (10 MHz). Moreover, the data can be processed by automated techniques
- Computers. The data from several channels can be fed directly to a high speed computer, either analog or digital, for on-line processing. The computer can separate parameters and calculate the variable of interest and significance, catalog the data, print summaries of the result, and store all data on tape for reference in future scans

Discontinuities Detectable by Eddy Current Inspection

Basically, any discontinuity that appreciably alters the normal flow of eddy currents can be detected by eddy current inspection. With encircling coil inspection of either solid cylinders or tubes, surface discontinuities having a combination of predominantly longitudinal and radial dimensional components are readily detected. When discontinuities of the same size are located beneath the surface of the part being inspected at progressively greater depths, they become increasingly difficult to detect, and can be detected at depths greater than 13 mm (½ in.) only with special equipment designed for this purpose.

Conversely, laminar discontinuities such as those in welded tubes might not alter the flow of the eddy currents enough to be detected unless the discontinuity breaks either the outside or inside surfaces, or unless it produces a discontinuity in the weld from upturned fibers caused by extrusion during welding. A similar difficulty could arise in trying to detect a thin planar discontinuity that is oriented substantially perpendicular to the axis of the cylinder.

Regardless of the limitations, a majority of objectionable discontinuities can be detected by eddy current inspection at high speed and at low cost. Some of the discontinuities that are readily detected are seams, laps, cracks, slivers, scabs, pits, slugs, open welds, miswelds, misaligned welds, black and gray oxide weld penetrators, pinholes, hook cracks, and surface cracks.

Reference Samples. A basic requirement for eddy current inspection is a reliable, consistent means to set tester sensitivity to the proper level each time it is used. A standard reference sample must be provided for this purpose. Without this capability, eddy current inspection is of little value. In selecting a standard reference sample, the usual procedure is to select a sample of product that can be run through the inspection system without producing appreciable indications from the tester. Several samples might have to be run before a suitable one is found; the suitable one then has reference discontinuities fabricated into it.

The type of reference discontinuities that must be used for a particular application are specified (for example, by ASTM and API). In selecting reference discontinuities, some of the major considerations are: (a) they must meet the required specification; (b) they should be easy to fabricate; (c) they should be reproducible; (d) they should be producible in precisely graduated sizes; and, (e) they should produce an indication on the eddy current tester that closely resembles those produced by the natural discontinuities.

Figure 27 shows several discontinuities that have been used for reference standards, these include a filed transverse notch; milled or electrical discharge machined longitudinal and transverse notches; and, drilled holes.

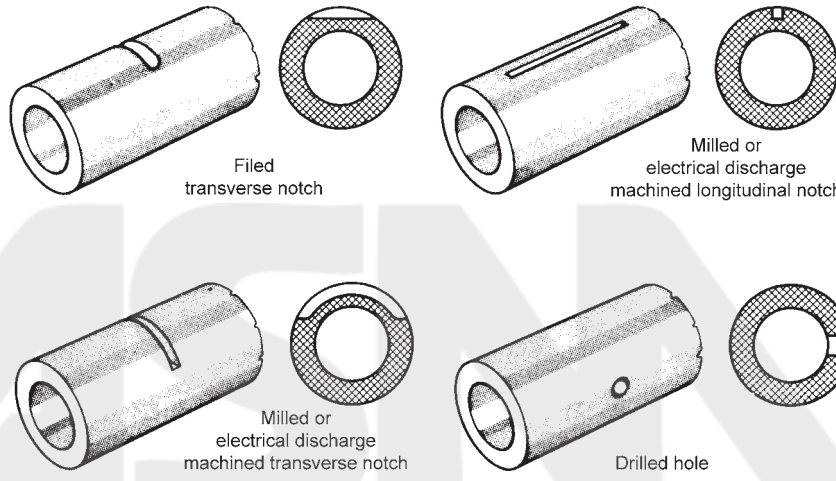


Fig. 27 Several fabricated discontinuities used as reference standards in eddy current inspection. ASTM standards for eddy current testing include E 215 (aluminum alloy tube), E 376 (measurement of coating thickness), E 243 (copper and copper alloy tube), E 566 (ferrous metal sorting), E 571 (nickel and nickel alloy tube), E 690 (nonmagnetic heat-exchanger tubes), E 426 (stainless steel tube), and E 309 (steel tube). Source: Ref 3

ACKNOWLEDGMENT

This chapter was adapted from Liquid-Penetrant Inspection, Magnetic-Particle Inspection, and Eddy-Current Inspection all in *Metals Handbook Desk Edition*, Second Edition, 1998.

REFERENCES

1. Liquid-Penetrant Inspection, *Metals Handbook Desk Edition*, 2nd ed., ASM International, 1998, p 1260–1267
2. Magnetic-Particle Inspection, *Metals Handbook Desk Edition*, 2nd ed., ASM International, 1998, p 1267–1273
3. Eddy-Current Inspection, *Metals Handbook Desk Edition*, 2nd ed., ASM International, 1998, p 1275–1281

SELECTED REFERENCES

- C. Hellier, *Handbook of Nondestructive Evaluation*, McGraw-Hill, 2000
- *Nondestructive Testing and Quality Control*, Vol 17, *ASM Handbook*, ASM International, 1989

CHAPTER **10**

Radiographic Inspection

RADIOGRAPHY is a nondestructive inspection method that is based on differential absorption of penetrating radiation—either electromagnetic radiation of very short wavelength or particulate radiation—by the part or test piece (object) being inspected. Because of differences in density and variations in thickness of the part, or differences in absorption characteristics caused by variations in composition, different portions of a test piece absorb different amounts of penetrating radiation. Unabsorbed radiation passing through the part can be recorded on film or photosensitive paper, viewed on a fluorescent screen, or monitored by various types of radiation detectors. The term *radiography* usually implies a radiographic process that produces a permanent image on film (conventional radiography) or paper (paper radiography or xeroradiography), although in a broad sense it refers to all forms of radiographic inspection. When inspection involves viewing of a real-time image on a fluorescent screen or image intensifier, the radiographic process is termed *real-time inspection*. When electronic, nonimaging instruments are used to measure the intensity of radiation, the process is termed *radiation gaging*. Tomography, a radiation inspection method adapted from the medical computerized axial tomography CAT scanner, provides a cross-sectional view of an inspection object. All the previous terms are used mainly in connection with inspection that involves penetrating electromagnetic radiation in the form of x-rays or gamma rays (also known as γ-rays). *Neutron radiography* refers to radiographic inspection using neutrons rather than electromagnetic radiation. This chapter discusses radiography methods using x-rays, gamma rays, and neutrons.

In conventional radiography, an object is placed in a beam of x-rays and the portion of the radiation that is not absorbed by the object impinges on a detector such as film. The unabsorbed radiation exposes the film emulsion, similar to the way that light exposes film in photography. Development of the film produces an image that is a two-dimensional *shadow pic-*

ture of the object. Variations in density, thickness, and composition of the object being inspected cause variations in the intensity of the unabsorbed radiation and appear as variations in photographic density (shades of gray) in the developed film. Evaluation of the radiograph is based on a comparison of the differences in photographic density with known characteristics of the object itself or with standards derived from radiographs of similar objects of acceptable quality.

Uses of Radiography

Radiography is used to detect features of a component or assembly that exhibit differences in thickness or physical density compared with surrounding material. Large differences are more easily detected than small ones. In general, radiography can detect only those features that have a reasonable thickness or radiation path length in a direction parallel to the radiation beam. This means that the ability of the process to detect planar discontinuities such as cracks depends on proper orientation of the test piece during inspection. Discontinuities such as voids and inclusions, which have measurable thickness in all directions, can be detected as long as they are not too small in relation to section thickness. In general, features that exhibit differences in absorption of a few percent compared with the surrounding material can be detected.

Applicability. Radiographic inspection is used extensively on castings and weldments, particularly where there is a critical need to ensure freedom from internal flaws. For instance, radiography often is specified for inspection of thick wall castings and weldments for steam power equipment, boiler and turbine components and assemblies, and other high pressure systems. Radiography also can be used on forgings and mechanical assemblies. When used with mechanical assemblies, radiography provides a unique nondestructive testing (NDT) capability of inspecting for condition and proper placement of components. Certain special devices are more satisfactorily inspected by radiography than by other methods. For instance, radiography is well suited to the inspection of semiconductor devices for cracks, broken wires, unsoldered connections, foreign material, and misplaced components, whereas other methods are limited in ability to inspect semiconductor devices.

Sensitivity of x-ray radiography, real-time x-ray methods, and gamma ray radiography to various types of flaws depends on many factors, including type of material, type of flaw, and product form. (Type of material in this context is usually expressed in terms of atomic number, for instance, metals having low atomic numbers are classified as light metals and those having high atomic numbers as heavy metals.) Table 1 indicates the general degrees of suitability of the three main radiographic methods for detection of discontinuities in various product forms and applications. In some instances, radiography cannot be used even though it appears

Table 1 Comparison of suitabilities of three radiographic methods for inspection of light and heavy metals

	Suitability for light metals(a)			Suitability for heavy metals(a)		
Inspection application	X-ray	Real-time radiography(b)	Gamma ray	X-ray	Real-time radiography(b)	Gamma ray
General						
Surface cracks(c)	F(d)	F(d)	F(d)	F(d)	F(d)	F(d)
Internal cracks	F(d)	F(d)	F(d)	F(d)	F(d)	F(d)
Voids	G	G	G	G	G	G
Thickness	F	F	F	F	F	F
Metallurgical variations	F	F	F	F	F	F
Sheet and plate						
Thickness	G(e)	G(e)	G(e)	G(e)	U	G(e)
Laminations	U	U	U	U	U	U
Voids	G	G	G	G	G	G
Bar and tube						
Seams	P	P	P	P	P	P
Pipe	G	G	G	G	F	F
Cupping	G	G	G	G	F	F
Inclusions	F	F	F	F	F	F
Castings						
Cold shuts	G	G	G	G	G	G
Surface cracks	F(d)	F(d)	F(d)	F(d)	F(d)	F(d)
Internal shrinkage	G	G	G	G	G	G
Voids, pores	G	G	G	G	G	G
Core shift	G	G	G	G	G	G
Forgings						
Laps	P(d)	P(d)	P(d)	P(d)	U	U
Inclusions	F	F	F	F	F	U
Internal bursts	G	G	G	F	F	G
Internal flakes	P(d)	P(d)	U	P(d)	P(d)	U
Cracks and tears	F(d)	F(d)	F(d)	F(d)	F(d)	F(d)
Welds						
Shrinkage cracks	G(d)	G(d)	G(d)	G(d)	G(d)	G(d)
Slag inclusions	G	G	G	G	G	G
Incomplete fusion	G	G	G	G	G	G
Pores	G	G	G	G	F	G
Incomplete penetration	G	G	G	G	G	G
Processing						
Heat treating cracks	U	F	U	P	P	U
Grinding cracks	U	F	U	U	U	U
Service						
Fatigue and heat cracks	F(d)	F(d)	P(d)	P	P	P
Stress corrosion	F	F	P	F	F	P
Blistering	P	P	P	P	P	P
Thinning	F	F	F	F	F	F
Corrosion pits	F	F	P	G	G	P

(a) G, good; F, fair: P, poor; U, unsatisfactory. (b) Real-time radiography offers the advantage that the part can be manipulated to present the best view—for example, align a crack. Also, when microfocus, magnification methods are used, real-time radiography presents excellent resolution and contrast. (c) Includes only visible cracks. Minute surface cracks normally are undetectable by radiographic inspection methods. (d) Radiation beam must be parallel to the cracks, laps, or flakes. (e) When calibrated using special thickness gages. Source: Ref 1

suitable from Table 1, because the part is accessible from one side only. Both sides must be accessible for radiography.

Radiography can be used to inspect most types of solid material, with the possible exception of assemblies containing materials of very high or very low density. Neutron radiography, however, often can be used in

such instances. Both ferrous and nonferrous alloys can be radiographed, as can nonmetallic materials and composites.

Limitations. Compared with other nondestructive methods of inspection, radiography is expensive. Relatively large capital costs and space allocations are required for a radiographic laboratory or a real time inspection station. Conversely, when portable x-ray or gamma ray sources are used, capital costs can be relatively low. Operating costs can be high; a large percentage of the total inspection time is spent in setting up for radiography. With real-time radiography, operating costs usually are much lower, because setup times are shorter and there are no extra costs for x-ray film and processing.

Field inspection of thick sections is a time-consuming process. Portable x-ray sources generally emit relatively low energy radiation, up to approximately 400 keV, and also are limited as to the intensity of radiation output. These characteristics of portable sources combine to limit x-radiography in the field to sections having absorption equivalent to that of approximately 75 mm (3 in.) of steel. Radioactive sources also are limited in the thickness that can be inspected, primarily because high activity sources require heavy shielding for protection of personnel. This limits field usage to sources of lower activity that can be transported in relatively lightweight containers. Because portable x-ray and gamma ray sources are limited in effective radiation output, exposure times usually are long for thick sections. Recent developments, such as a portable linear accelerator, can speed up and increase the penetrating power of field radiographic methods.

Certain types of flaws are difficult to detect by radiography. Laminar defects such as cracks present problems unless they are essentially parallel to the radiation beam. Tight, meandering cracks in thick sections usually cannot be detected even when properly oriented. Minute discontinuities such as inclusions in wrought material, flakes, microporosity, and microfissures cannot be detected unless they are sufficiently segregated to yield a detectable gross effect. Laminations normally are not detectable by radiography because of their unfavorable orientation, usually parallel to the surface. Laminations seldom yield differences in absorption that enable laminated areas to be distinguished from lamination-free areas.

Principles of Radiography

There are three basic elements that combine to produce a radiograph: a radiation source or probing medium; the test piece or object being evaluated; and, a recording medium (usually film), as shown schematically in Fig. 1. The test piece is a plate of uniform thickness containing an internal flaw that has absorption characteristics different from those of the surrounding material. Radiation from the source is absorbed by the test piece as the radiation passes through it; the flaw and surrounding material ab-

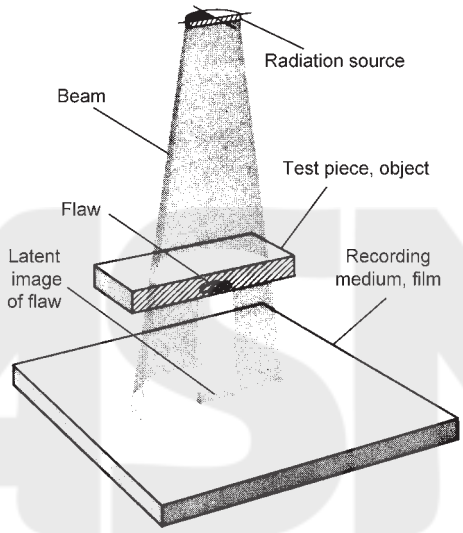


Fig. 1 Diagram of the basic elements of a radiographic system, showing method of detecting and recording an internal flaw in a plate of uniform thickness. Source: Ref 1

sorb different amounts of radiation. Thus, the amount of radiation that reaches the film in the area beneath the flaw is different from the amount that impinges on adjacent areas. This produces on the film a latent image of the flaw that, when the film is developed, can be seen as a *shadow* of different photographic density from that of the image of the surrounding material.

Geometric Factors In Radiography. Because a radiograph is a two-dimensional representation of a three-dimensional object, the radiographic images of most test pieces are somewhat distorted in size and shape.

In conventional radiography, the position of a flaw within the volume of a test piece cannot be determined exactly with a single radiograph; depth in the direction of the radiation beam cannot be determined exactly. Conclusions regarding depth sometimes can be drawn from the sharpness of the flaw image. Images of flaws close to the detector tend to appear sharper than images of flaws near the source side of the object. However, techniques such as stereoradiography, tomography, triangulation, or simply making two or more exposures, with the radiation beam being directed at the test piece from a different angle for each exposure, can be used to locate flaws more exactly within the test-piece volume.

Sources of Radiation

Two types of electromagnetic radiation are used in radiographic inspection: x-rays and γ-rays. X-rays and γ-rays differ in their wavelengths from other types of electromagnetic radiation such as visible light, microwaves,

and radio waves; although there is not always a distinct transition from one type of electromagnetic radiation to another (Fig. 2). Only x-rays and γ-rays, because of their relatively short wavelengths (high energies), have the capability of penetrating opaque materials to reveal internal flaws.

X-rays and γ-rays are physically indistinguishable; they differ only in the manner in which they are produced. X-rays result from the interaction between a rapidly moving stream of electrons and atoms in a solid target material, while γ-rays are emitted during the radioactive decay of unstable atomic nuclei.

The amount of exposure from x-rays or γ-rays is measured in roentgens (R), where 1 R is the amount of radiation exposure that produces one electrostatic unit (3.33564×10^{-10} C) of charge from 1.293 mg (45.61×10^{-6} oz) of air. The intensity of an x-ray or γ-ray radiation is measured in roentgens per unit time.

Although the intensity of x-ray or γ-ray radiation is measured in the same units, the strengths of x-ray and γ-ray sources are usually given in different units. The strength of an x-ray source is typically given in roentgens per minute at one meter (RMM) from the source or in some other suitable combination of time or distance units (such as roentgens per hour at one meter, or RHM). The strength of a γ-ray source is usually given in terms of the radioactive decay rate, which has the traditional unit of a Curie (1 Ci = 37×10^9 disintegrations per second). The corresponding unit in the Système International d'Unités (SI) system is a gigabecquerel (1 GBq = 1×10^9 disintegrations per second).

The spectrum of radiation is often expressed in terms of photon energy rather than as a wavelength. Photon energy is measured in electron volts (eV), with 1 eV being the energy imparted to an electron by an accelerating potential of 1 V. The radiation spectrum in terms of both wavelength and photon energy is shown in Fig. 2.

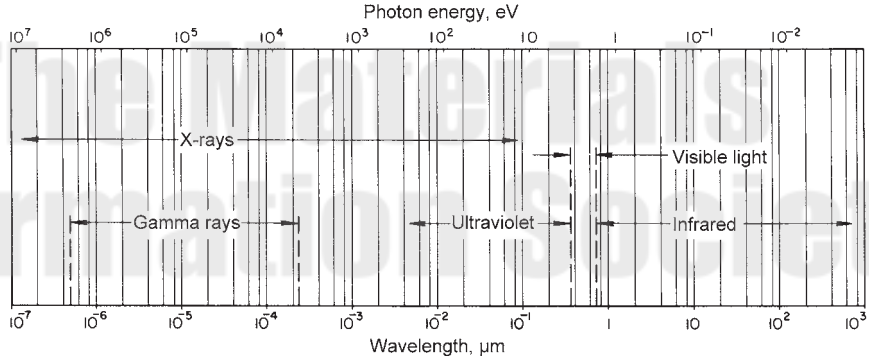


Fig. 2 Schematic representation of the portion of the electromagnetic spectrum that includes x-rays, gamma rays, ultraviolet and visible light, and infrared radiation, showing their relationship with wave length and photon energy. Source: Ref 1

Production of X-Rays. When x-rays are produced from the collision of fast moving electrons with a target material, two types of x-rays are generated. The first type of x-ray is generated when the electrons are rapidly decelerated during collisions with atoms in the target material. These x-rays have a broad spectrum of many wavelengths (or energies) and are referred to as *continuous x-rays* or by the German word *bremsstrahlung*, which means braking radiation. The second type of x-ray occurs when the collision of an electron with an atom of the target material causes a transition of an orbital electron in the atom, thus leaving the atom in an excited state. When the orbital electrons in the excited atom rearrange themselves, x-rays are emitted that have specific wavelengths (or energies) characteristic of the particular electron rearrangements taking place. These characteristic x-rays usually have much higher intensities than the background of bremsstrahlung having the same wavelengths.

Production of γ-Rays. Gamma rays are generated during the radioactive decay of both naturally occurring and artificially produced unstable isotopes. In all respects other than their origin, γ-rays and x-rays are identical. Unlike the broad spectrum radiation produced by x-ray sources, γ-ray sources emit one or more discrete wavelengths of radiation, each having its own characteristic photon energy (or wavelength).

Many of the elements in the periodic table have either naturally occurring radioactive isotopes or isotopes that can be made radioactive by irradiation with a stream of neutrons in the core of the nuclear reactor. However, only certain isotopes are extensively used for radiography, as shown in Table 2.

X-Ray Tubes

X-ray tubes are electronic devices that convert electrical energy into x-rays. Typically, an x-ray tube consists of a cathode structure containing a filament and an anode structure containing a target, all within an evacuated chamber or envelope, as illustrated in Fig. 3. A low-voltage power supply, usually controlled by a rheostat, generates the electric current that heats the filament to incandescence. This incandescence of the filament produces an electron cloud, which is directed to the anode by a focusing

Table 2 Characteristics of γ-ray sources used in industrial radiography

γ-ray source	Half-life	Photon energy, MeV	Radiation output(a), RHM/Ci	Penetrating power, mm (in.) of steel
Thulium-170	128 d	0.054 and 0.084(b)	0.003	13 (0.5)
Iridium-192	74 d	12 rays from 0.21–0.61	0.48	75 (3)
Cesium-137	33 yr	0.66	0.32	75 (3)
Cobalt-60	5.3 yr	1.17 and 1.33	1.3	230 (9)

(a) Output for typical unshielded, encapsulated sources: RHM/Ci, roentgens per hour at 1 m per Curie. (b) Against strong background of higher MeV radiation. Source: Ref 1

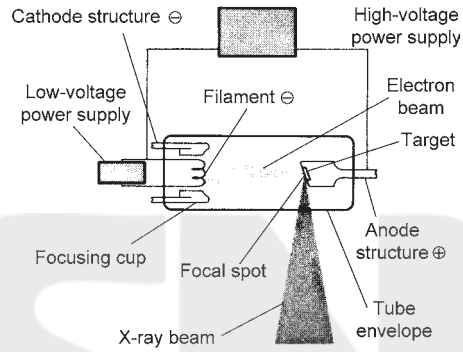


Fig. 3 Schematic diagram of the principal components of an x-ray unit. Source: Ref 1

system and accelerated to the anode by the high voltage applied between the cathode and the anode. Depending on the size of the focal spot achieved, x-ray tubes are sometimes classified into three groups:

- Conventional x-ray tubes with focal spot sizes between 2 by 2 mm (0.08 by 0.08 in.) and 5 by 5 mm (0.2 by 0.2 in.)
- Minifocus tubes with focal spot sizes in the range of 0.2 mm (0.008 in.) and 0.8 mm (0.03 in.)
- Microfocus tubes with focal spot sizes in the range of 0.005 mm (0.0002 in.) and 0.05 mm (0.002 in.)

There are three important electrical characteristics of x-ray tubes:

- The filament current that controls the filament temperature and in turn the quantity of electrons that are emitted
- The tube voltage or anode-to-cathode potential that controls the energy of impinging electrons and the energy or penetrating power of the x-ray beam
- The tube current that is directly related to the filament temperature and is usually referred to as the milliamperage of the tube

The strength or radiation output of the beam is approximately proportional to milliamperage, which is used as one of the variables in exposure calculations. This radiation output or R output is usually expressed in roentgens per minute (or hour) at 1 m.

When the accelerated electrons impinge on the target immediately beneath the focal spot, the electrons are slowed and absorbed, and both bremsstrahlung and characteristic x-rays are produced. Most of the energy in the impinging electron beam is transformed into heat, which must be dissipated. Severe restrictions are imposed on the design and selection of materials for the anode and target to ensure that structural damage from overheating does not prematurely destroy the target. Anode heating also

limits the size of the focal spot. Because smaller focal spots produce sharper radiographic images, the design of the anode and target represents a compromise between maximum radiographic definition and maximum target life. In many x-ray tubes, a long, narrow, actual focal spot is projected as a roughly square effective focal spot by inclining the anode face at a small angle (usually about 20°) to the centerline of the x-ray beam, as shown in Fig. 4.

Tube Design and Materials. The cathode structure in a conventional x-ray tube incorporates a filament and a focusing cup, which surrounds the filament. The focusing cup, usually made of pure iron or pure nickel, functions as an electrostatic lens whose purpose is to direct the electron beam toward the anode. The filament, usually a coil of tungsten wire, is heated to incandescence by an electric current produced by a relatively low voltage, similar to the operation of an ordinary incandescent light bulb. At incandescence, the filament emits electrons that are accelerated across the evacuated space between the cathode and the anode. The driving force for acceleration is a high electrical potential (voltage) between anode and cathode, which is applied during exposure.

The anode usually consists of a button of the target material embedded in a mass of copper that absorbs much of the heat generated by electron collisions with the target. Tungsten is the preferred material for traditional x-ray tubes used in radiography because its high atomic number makes it an efficient emitter of x-rays and because its high melting point enables it to withstand the high temperatures of operation. Gold and platinum are also used in x-ray tubes for radiography, but targets made of these metals must be more effectively cooled than targets made of tungsten. Other materials are used, particularly at low energies, to take advantage of their characteristic radiation. Most high intensity x-ray tubes have forced liquid cooling to dissipate the large amounts of anode heat generated during operation.

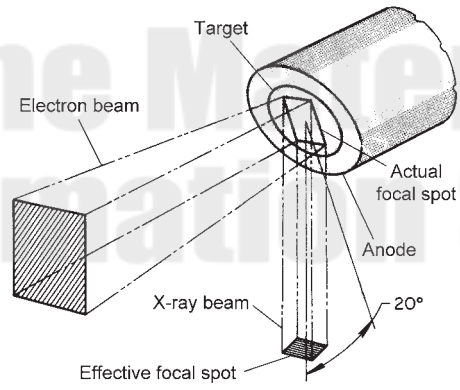


Fig. 4 Schematic diagram of the actual and effective focal spots of an anode that is inclined at 20° to the centerline of the x-ray beam. Source: Ref 1

Tube envelopes are constructed of glass, ceramic materials or metals, or combinations of these materials. X-ray tubes are inserted into metallic housings that contain an insulating medium such as transformer oil or an insulating gas. The main purpose of the insulated housing is to provide protection from high voltage electrical shock. Housings usually contain quick disconnects for electrical cables from the high-voltage power supply or transformer. On self-contained units, most of which are portable, both the x-ray tube and the high voltage transformer are contained in a single housing, and no high-voltage cables are used.

Microfocus X-Ray Tubes. Developments in vacuum technology and manufacturing processes have led to the design and manufacture of microfocus x-ray systems. These systems are available with voltages varying from 10 to 360 kV at beam currents from 0.01 to 2 mA. To avoid excessive pitting of the target, the beam current is varied according to the desired focal spot size and/or kilovolt level.

Microfocus x-ray systems having focal spots that approach a point source are useful in obtaining very high resolution images. A radiographic definition of 20 line pairs per millimeter, or a spatial resolution of 0.002 in., using real-time radiography has been achieved with microfocus x-ray sources. This high degree of radiographic definition is accomplished by image enlargement, which allows the imaging of small details.

Microfocus x-ray systems have found considerable use in the inspection of integrated circuits and other miniature electronic components. Microfocus x-ray systems with specially designed anodes as small as 13 mm (0.5 in.) in diameter and several inches long also enable an x-ray source to be placed inside otherwise inaccessible areas, such as aircraft structures and piping. The imaging medium is placed on the exterior, and this allows for the single wall inspection of otherwise uninspectable critical components. Because of the small focal spot, the source can be close to the test area with minimal geometric unsharpness.

X-Ray Spectrum. The output of a radiographic x-ray tube is not a single wavelength beam, but rather a spectrum of wavelengths somewhat analogous to white light. The lower limit of wavelengths, λ_{min}, in nanometers, at which there is an abrupt ending of the spectrum, is inversely proportional to tube voltage, V. This corresponds to an upper limit on photon energy, E_{max}, which is proportional to the tube voltage, V:

$$E_{max} = aV$$

where a = 1 eV/volt.

Figure 5 illustrates the effect of variations in tube voltage and tube current on photon energy and the intensity (number of photons). As shown in Fig. 5(a), increasing the tube voltage increases the intensity of radiation and adds higher energy photons to the spectrum (crosshatched area in Fig. 5a). Conversely, as shown in Fig. 5(b), increasing the tube current increases the intensity of radiation but does not affect the energy distribution.

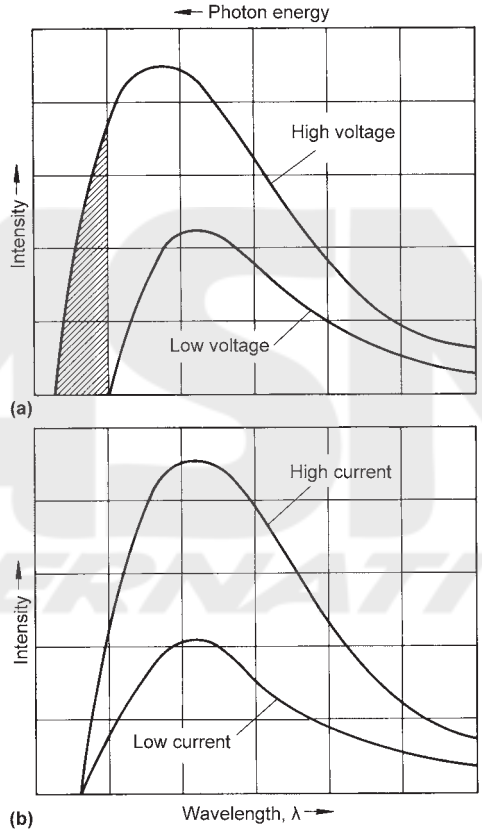


Fig. 5 Effect of (a) tube voltage and (b) tube current on the variation of intensity with wavelength for the bremsstrahlung spectrum of an x-ray tube. See text for discussion. Source: Ref 1

The energy of x-rays determines the penetration capability. Table 3 gives penetrating capabilities of x-ray beams of various energy levels expressed as the range of steel thickness that can be satisfactorily inspected. The maximum values in this table represent thicknesses of steel that can be routinely inspected using exposures of several minutes' duration and with medium speed film. Thicker sections can be inspected for each x-ray energy value by using faster films and long exposure times, but for routine work the use of higher energy x-rays is more practical. Sections thinner than minimum thicknesses shown in Table 3 can easily be penetrated, but radiographic contrast may not be optimum.

Attenuation of Electromagnetic Radiation

X-rays and gamma rays interact with any substance, even gases such as air, as the rays pass through the substance. It is this interaction that enables parts to be inspected by differential attenuation of radiation, and that enables differences in the intensity of radiation to be detected and recorded.

Table 3 Penetrating capabilities of conventional x-ray tubes and high energy sources

	Penetration range for steel	
Maximum accelerating potential	**mm**	**in.**
X-ray tubes		
150 kV	Up to 16	Up to ⅝
250 kV	Up to 38	Up to 1½
400 kV	Up to 64	Up to 2½
1000 kV (1 MV)	6.4 to 89	¼ to 3½
High-energy sources		
2.0 MeV	6.4 to 250	¼ to 10
4.5 MeV	25 to 305	1 to 12
7.5 MeV	57 to 460	2¼ to 18
20.0 MeV	75 to 610	3 to 24

Source: Ref 1

Both these effects are essential to the radiographic process. Attenuation characteristics of materials vary with type, intensity, and energy of the radiation, and with density and atomic structure of the material.

The attenuation of electromagnetic radiation is a complex process. The intensity of radiation varies exponentially with the thickness of homogeneous material through which it passes. This behavior is expressed as:

$$I = I_o \exp(-\mu t)$$

where I is the intensity of the emergent radiation, I_0 is the initial intensity, t is the thickness of homogeneous material, and μ is a characteristic of the material known as the linear absorption coefficient. The coefficient μ is constant for a given situation but varies with the material and with the photon energy of the radiation. The units of μ are reciprocal length (for instance, cm^{-1}). The absorption coefficient of a material is sometimes expressed as a mass-absorption coefficient (μ/ρ), where ρ is the density of the material.

There are three primary attenuation processes: photoelectric effect, Compton scattering, and pair production.

Photoelectric effect is an interaction with orbital electrons in which a photon of electromagnetic radiation is consumed in breaking the bond between an orbital electron and its atom. Energy in excess of the bond strength imparts kinetic energy to the electron.

The photoelectric effect generally decreases with increasing photon energy (E) as $E^{-3.5}$. For elements of low atomic number, the photoelectric effect is negligible at photon energies exceeding about 100 keV. However, the photoelectric effect varies with the fourth to fifth power of atomic number; thus, for elements of high atomic number, the effect accounts for an appreciable portion of total absorption at photon energies up to about 2 MeV.

Compton scattering is a form of direct interaction between an incident photon and an orbital electron in which the electron is ejected from the

atom and only a portion of the kinetic energy of the photon is consumed. The photon is scattered incoherently, emerging in a direction that is different from the direction of incident radiation and emerging with reduced energy. The relationship of the intensity of the scattered beam to the intensity of the incident beam, scattering angle, and photon energy in the incident beam is complex, yet is amenable to theoretical evaluation. Compton scattering varies directly with atomic number of the scattering element and approximately inversely with photon energy in the energy range that is of major interest.

Pair production is an absorption process that creates two 0.5 MeV photons of scattered radiation for each photon of high energy incident radiation consumed; a small amount of scattered radiation of lower energy also accompanies pair production. Pair production is more important for heavier elements; the effect varies with atomic number, *Z*, approximately as $Z(Z + 1)$. The effect also varies approximately logarithmically with photon energy.

In pair production, a photon of incident electromagnetic radiation is consumed in creating an electron-positron pair that then is ejected from an atom. This effect is possible only at photon energies exceeding 1.02 MeV because, according to the theory of relativity, 0.51 MeV is consumed in the creation of the mass of each particle, electron, or positron. Any energy of the incident photon exceeding 1.02 MeV imparts kinetic energy to the pair of particles.

Radiographic Equivalence. The absorption of x-rays and gamma rays by various materials becomes less dependent on composition as radiation energy increases. For instance, at 150 kV, 25 mm (1 in.) of lead is equivalent to 350 mm (14 in.) of steel, but at 1000 kV, 25 mm of lead is equivalent to only 125 mm (5 in.) of steel. Approximate radiographic absorption equivalence factors for several metals are given in Table 4. When exposure charts are only available for certain common materials, such as steel

Table 4 Approximate radiographic absorption equivalence for various metals

	X-rays, kV					X-rays, MeV			Gamma rays			
Material	**50**	**100**	**150**	**220**	**400**	**1**	**2**	**4-25**	**Ir-192**	**Cs-137**	**Co-60**	**Ra**
Magnesium	0.6	0.6	0.05	0.08	...	...	...	...	...	...	...	...
Aluminum	1.0	1.0	0.12	0.18	...	...	...	...	0.35	0.35	0.35	0.40
Aluminum alloy 2024	2.2	1.6	0.16	0.22	...	...	...	...	0.35	0.35	0.35	...
Titanium	...	...	0.45	0.35	...	...	...	...	...	...	...	...
Steel	...	12.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
18-8 stainless steel	...	12.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Copper	...	18.0	1.6	1.4	1.4	...	...	1.3	1.1	1.1	1.1	1.1
Zinc	...	...	1.4	1.3	1.3	...	...	1.2	1.1	1.0	1.0	1.0
Brass(a)	...	...	1.4	1.3	1.3	1.2	1.2	1.2	1.1	1.1	1.1	1.1
Inconel alloys	...	16.0	1.4	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
Zirconium	...	...	2.3	2.0	...	1.0	...	...	...	...	...	...
Lead	...	...	14.0	12.0	...	5.0	2.5	3.0	4.0	3.2	2.3	2.0
Uranium	...	...	...	25.0	...	...	...	3.9	12.6	5.6	3.4	...

Source: (a) Containing no tin or lead; absorption equivalence is greater than these values when either element is present. Source: Ref 1

or aluminum, exposure times for other materials can be estimated by determining the exposure time for an equal thickness of a common material from the chart, then multiplying by the radiographic equivalence factor.

Principles of Shadow Formation

The image formed on a radiograph is similar to the shadow cast on a screen by an opaque object placed in a beam of light. Although radiation used in radiography penetrates an opaque object whereas light does not, the geometric laws of shadow formation are basically the same. X-rays, gamma rays, and light all travel in straight lines. Straight line propagation is the chief characteristic of radiation that permits formation of a sharply discernible shadow. The geometric relationships of source, object, and screen to each other determine the three main characteristics of the shadow: the degrees of enlargement, distortion, and unsharpness (see Fig. 6).

Enlargement. The shadow of the object (test piece) is always farther from the source than the object itself. Thus, as illustrated for a point source in Fig. 6(a), dimensions of the shadow are always greater than corresponding dimensions of the object. Mathematically, the size of the image or degree of enlargement may be calculated from the relationship:

$$M = S_i/S_o = L_i/L_o$$

where M is the degree of enlargement (magnification), S_i is the size of the image, S_o is the size of the object, L_i is the source-to-image distance, and L_o is the source-to-object distance.

With very small focal spots, large values of geometric magnification can be used effectively. Values of 6 to 20 are common; magnification values as high as 100 can be used. Focal spots in microfocal x-ray equipment range from 5 to 20 μm (0.0002 to 0.0008 in.). In addition to increased image size, magnification systems also offer improved contrast because radiation scattered in the object does not reach the detector.

Distortion. As long as the plane of a two-dimensional object and the plane of the recording surface are parallel to each other, the image of that object plane will be undistorted regardless of the angle at which the beam of radiation impinges on the object. Also, the degree of enlargement for different points in a given object plane is constant because the ratio L_i/L_o is invariant. However, as shown in Fig. 6(b), if the plane of the object and the plane of the recording surface are not parallel, the image will be distorted. For objects of appreciable thickness, the magnification for different object planes will vary because L_o varies.

Geometric Unsharpness. In reality, most radiation sources are too large to be approximated by a point. Most conventional x-ray tubes have focal spots several millimeters in size. Even high energy sources have focal spots of appreciable size, although seldom exceeding 2 mm (0.08

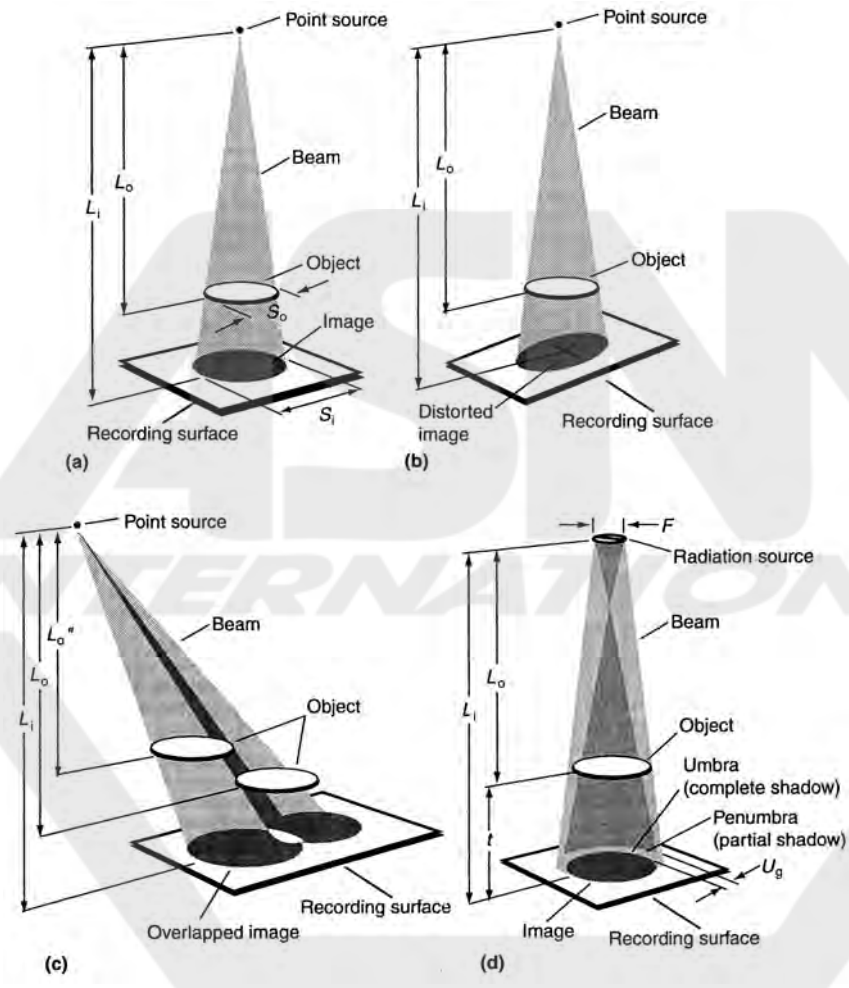


Fig. 6 Schematic representation of the effect of geometric relationships on radiographic image from point sources and actual radiation source. (a) Image size. (b) Image distortion. (c) Image overlap for point sources of radiation. (d) Degree of image unsharpness from an actual radiation source. See text for discussion. L_o = source-to-object distance; L_i = source-to-image distance; S_o = size of object; S_i = size of image; U_g = geometric unsharpness; F = size of focal spot; t = object-to-image distance. Source: Ref 1

in.) in diameter. Gamma ray sources vary widely in size, depending on source strength and specific activity, but seldom are less than about 2.5 mm (0.1 in.) in diameter.

Geometric unsharpness is one of several unsharpness factors, and, at low and medium x-ray energies, is usually the largest contributor to maximum unsharpness. Neglecting the distance between the actual surface of the recording medium and the adjacent (facing) surface of the test piece, which usually is quite small in relation to test piece thickness, the geometric unsharpness can be calculated for any source size, and can be expressed

as a series of straight-line plots relating geometric unsharpness, U_g, to test-piece thickness, t, for various values of source-to-object distance, L_o. A typical series is shown in Fig. 7 for a 5 mm (0.2 in.) diameter source. It is helpful to prepare graphs like the one in Fig. 7 for each source size used.

Image Conversion

The most important process in radiography is the conversion of radiation into a form suitable for observation or further signals processing. After penetrating the test piece, the x-rays or γ-rays pass through a medium on the imaging surface. This medium may be a recording medium such as film, or a medium that responds to the intensity of the radiation, such as fluorescent screen or a scintillation crystal in a discrete detector. These two types of media provide the images for subsequent observation.

Radiography can also utilize radiographic screens, which are pressed into intimate contact with the imaging medium. Radiographic screens include:

- Metallic screens placed over film, paper, or the screens used in real-time systems
- Fluorescent screens placed over film or photographic paper
- Fluorometallic screens placed over film

These screens are used to improve radiographic contrast by intensifying the conversion of radiation and by filtering the lower energy radiation produced by scattering.

Permanent images are recorded on x-ray film, radiographic paper, or electrostatically sensitive paper such as is used in the xeroradiographic process and are called radiographs. Real-time images, such as those presented on a fluorescent screen, image amplifier, or television monitor, dif-

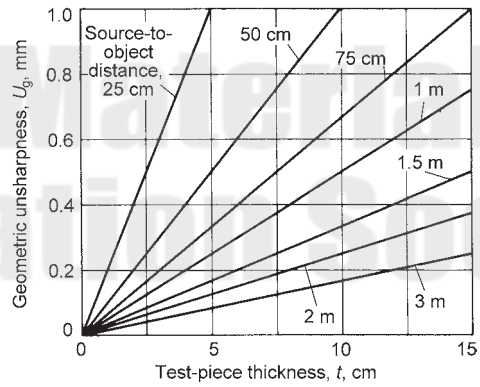


Fig. 7 Relation of geometric unsharpness to testpiece thickness for various source-to-object distances when the source is 5 mm in diameter. Source: Ref 1

fer in appearance from those on radiographs; records of these images may be made by photography or video recording. If the information is sensed or recorded using radiation measuring instruments and does not appear as an image, the recording process is termed *radiation gaging*. X-ray film is used more extensively than all other recording media combined.

X-ray film is constructed of a thin, transparent plastic support called a film base, which usually is coated on both sides (but occasionally on one side only) with an emulsion consisting mainly of grains of silver salts that are embedded in gelatin (see Fig. 8). These salts are very sensitive to electromagnetic radiation, especially x-rays, gamma rays, and visible light. The film base, usually tinted blue, is approximately 0.18 mm (0.007 in.) thick. An adhesive undercoat fastens the emulsion to the film base. A very thin but tough coating of gelatin called a protective overcoat covers the emulsion to protect it against minor abrasion. The total thickness of the x-ray film is approximately 0.23 mm (0.009 in.), including film base, two emulsions, two adhesive undercoats, and two protective overcoats.

Radiographic Paper. Ordinary photographic paper can be used to record x-ray images, although its characteristics are not always satisfactory. Photographic paper has a low speed and the resulting image is low in contrast. However, photographic paper in various forms can be used effectively for some applications.

Radiographic paper can exhibit excellent sensitivity, which in many respects matches or exceeds that of fast direct exposure x-ray films. Radiographic paper does not match the sensitivities of slow x-ray films, but because of their speed, convenience, and low cost, radiographic papers are being used both for radiography of materials that do not require critical examination and for "in-process" control.

Xeroradiography (dry radiography) is a form of imaging that uses electrostatic principles for formation of a radiographic image. In film radiography, a latent image is formed in the emulsion of a film. In xeroradiography, the latent image is formed on a plate coated with a photoconductive layer of selenium. Before use, the plate is given an even charge of static electricity over the entire surface. As soon as the plate is charged, it be-

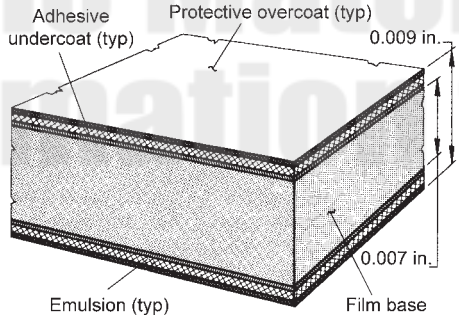


Fig. 8 Schematic cross-section of a typical x-ray film. Source: Ref 1

comes sensitive to light as well as to x-radiation and must be protected from light by a rigid holder similar to a film cassette. In practice, the holder is used for radiography as though it contained film. X-radiation will differentially discharge the plate according to the amount of radiation received by different areas. This forms an electrostatic latent image of the test piece on the plate.

Development of the exposed plate is done by subjecting the plate, in the absence of light, to a cloud of fine powder charged opposite to the electrostatic charges remaining on the plate. The charged powder is attracted to the residual charges on the plate. The visible radiographic image can be made permanent by placing a piece of specially treated paper over the plate and transferring the powder to the paper, which then is heated to fix the powder in place.

Selenium coated plates can be easily damaged by fingerprints, dirt, and abrasion. For this reason, automated equipment is used for charging and for development and image transfer to paper.

Fluorescent screens consist of crystals that emit light in proportion to the intensity of the impinging radiation. The real-time image can then be viewed directly with appropriate measures to protect the viewer from radiation or can be monitored by low level television camera tubes.

Direct viewing of fluoroscopic images is known as fluoroscopy. This method is the predecessor of the modern methods of real-time radiography but fluoroscopy is now largely obsolete. The main problem with fluoroscopy is the low level of light output from the fluorescent screen. This requires the suppression of background light and about 30 minutes for the viewer's eyes to become acclimated. Moreover, radiation safety dictates viewing through leaded glass or indirectly by mirrors. Because of these limitations, image intensities have been developed to improve safety and to amplify the images from fluorescent screens.

The modern development of low-level television camera tubes and low noise video circuitry also allows video monitoring of the dim images on fluorescent screens. The contrast sensitivity and the spatial resolution of fluorescent screen systems are comparable to those of image intensifiers, but the use of fluorescent screens is limited to lower radiation energies, below about 320 keV without intensifying screens and about 1 MeV with intensifying screens. Nevertheless, fluorescent screens can provide an unlimited field of view, while image intensifiers have a field of view limited to approximately 300 mm (12 in.). The dynamic range of systems with fluorescent screens can vary from 20 to 1 for raw images to 1000 to 1 with digital processing and a large number of frames averaged.

Image intensifier tubes are glass enclosed vacuum devices that convert a low-intensity x-ray image or a low-brightness, fluorescent screen image into a high-brightness, visible light image. Image intensification is achieved by a combination of electronic amplification and image minification. The image brightness at the output window of an image intensifier

tube is approximately 0.3×10^3 cd/m^2 (10^{-1} lambert), as compared to approximately 0.3 cd/m^2 (10^{-4} lambert) for a conventional fluoroscopic screen.

The early image intensifiers were originally developed for medical purposes and were limited to applications with low-energy radiation because of low detection efficiencies at high energies. Consequently, industrial radiography with these devices was restricted to aluminum, plastics, or thin sections of steel. By the mid-1970s, other technological developments led to further improvements such as high energy, x-ray sensitivity for image intensifiers, improved screen materials, digital video processing for image enhancement, and high definition imaging with microfocus x-ray generators.

The early image intensifiers were only suitable for medical applications and the inspection of light materials and thin sections of steel. The image quality was not sufficient for general use in radiography, and, image intensifiers had to be redesigned for industrial material testing. The modern image intensifier for industrial application is a very practical imaging device for radiographic inspection with radiation energies up to 10 MeV. With the image intensifier, a 2% difference in absorption can be routinely achieved in production inspection applications. The typical dynamic range of an image intensifier before image processing is about 2000 to 1.

Digital Radiography. Another method of radiographic imaging involves the formation of an image by scanning a linear array of discrete detectors along the object being irradiated. This method directly digitizes the radiometric output of the detectors and generates images in near real time. Direct digitization (as opposed to digitizing the output of a TV camera or image intensifier) enhances the signal-to-noise ratio and can result in a dynamic range up to 100,000 to 1. The large dynamic range of digital radiography allows the inspection of parts having a wide range of thicknesses and densities. Discrete detector arrangements also allow the reduction of secondary radiation from scattering by using a fan beam detector arrangement like that of computed tomography (CT) systems. In fact, industrial CT systems are used to obtain digital radiographs. The imaging performance with detector arrays is also comparable with that of computed tomography (Table 5).

The detectors used in digital radiography include scintillator photodetectors, phosphor photodetectors, photomultiplier tubes, and gas ionization detectors. Scintillator and phosphor photodetectors are compact and rugged, and they are used in flying spot and fan beam detector arrangements. Photomultiplier tubes are fragile and bulky, but do provide the capability of photon counting when signal levels are low. Gas ionization detectors have low detection efficiencies but better long-term stability than scintillator and phosphor photodetector arrays.

A typical phosphor photodetector array for the radiographic inspection of welds consists of 1024 pixel elements with 25 μm (1 mil) spacing, cov-

Table 5 Comparison of performance characteristics for film radiography, real-time radiography, and x-ray computed tomography

Performance characteristic	Film radiography	Real-time radiography(a)	Computed tomography or digital radiography(b)
Spatial resolution(c)	>5 line pairs/mm	~2.5 line pairs/mm	0.2–4.5 line pairs/mm
Absorption efficiencies, %			
Absorption efficiency (80 keV)	5	20	99
Absorption efficiency (420 keV)	2	8	95
Absorption efficiency (2 MeV)	0.5	2	80
Sources of noise	Scatter, poor photon statistics	Scatter, poor photon statistics	Minimal scatter
Dynamic range	200–1000	500–2000	Up to 1×10^6
Digital image processing	Poor, requires film scanner	Moderate to good; typically 8-bit data	Excellent; typically 16-bit data
Dimensioning capability	Moderate; affected by structure visibility and variable radiographic magnification	Moderate to poor; affected by structure visibility, resolution, variable radiographic magnification, and optical distortions	Excellent; affected by resolution, enhanced by low contrast detectability

Source: (a) General characteristics of real-time radiography with fluorescent screen-TV camera system or an image intensifier. (b) Digital radiographic imaging performance with discrete element detector arrays is comparable to computed tomography performance values. (c) Can be improved with microfocus x-ray source and geometric magnification. Source: Ref 1

ering 25 mm (1 in.) in length perpendicular weld seam. The linear photodiode array is covered with a fiberoptic faceplate and can be cooled in order to reduce noise. For the conversion of the x-rays to visible light, fluorescent screens are coupled to the array by means of the fiber optics. A linear collimator parallel to the array is arranged in front of the screen. The resolution perpendicular to the array is defined by the width of this slit and the speed of the manipulator. A second, single element detector is provided to detect instabilities of the x-ray beam. Using 100 kV radiation, a spatial resolution of 0.1 mm (0.004 in.) can be achieved with a scanning speed of 1 to 10 mm/s (0.04 to 0.4 in./s).

The data from the detector system are digitized and then stored in a fast dual ported memory. This permits quasi simultaneous access to the data during acquisition. Before the image is stored in the frame buffer and displayed on the monitor, simple preprocessing can be done, such as intensity correction of the x-ray tube by the data of the second detector and correction of the sensitivity for different array elements. If further image processing or automatic defect evaluation is required, the system can be equipped with fast image processing hardware. All standard devices for digital storage can be utilized.

Image Processing. Because real-time systems generally do not provide the same level of image quality and contrast as radiographic film, image processing is often used to enhance the images from image intensifiers, fluorescent screens, and detector arrays. With image processing, the video images from real-time systems can compete with the image quality of film radiography. Moreover, image processing also increases the dynamic

range of real-time systems beyond that of film, which typically has a dynamic range of about 1000 to 1.

Images can be processed in two ways: as an analog video signal and as a digitized signal. An example of analog processing is to shade the image after the signal leaves the camera. Shading compensates for irregularities in brightness across the video image due to thickness or density variations in the test piece. This increases the dynamic range (or latitude) of the system, which allows the inspection of parts having larger variations in thickness and density.

After analog image processing, the images can be digitized for further image enhancement. This digitization of the signal may involve some detector requirements. In all real-time radiological applications, the images have to be obtained at low dose rates (around 20 μR/s). However, in digital x-ray imaging, the most often used dose is around 1 mR, in order to reduce the signal fluctuations that would result from a weak x-ray flux. This means that only a short exposure time (of the order of a few milliseconds) and a frequency of several images can be used if kinetic (motion) blurring is to be avoided. The resulting requirements for the x-ray detector are:

- The capability of operating properly in a pulsed mode, which calls for a fast temporal response
- An excellent linearity, to allow the use of the simplest and most efficient form of signal processing
- A wide operating dynamic range in terms of dose output (around 2000 to 1)

Once the image from the video camera has been digitized in the image processor, a variety of processing techniques can be implemented. The image processing techniques may range from the relatively simple operation of frame integration to more complex operations such as automatic defect evaluation.

Radiation Gaging. Radiation measuring instruments do not produce images. The output from these instruments is a meter reading or a strip chart, which records the radiation transmitted through a test piece in terms of roentgens. Many of these instruments are routinely used to check areas surrounding a radiographic inspection site for excessive radiation.

Radiation gaging can be applied to certain automated processes, such as thickness gaging of materials or determination of liquid levels in sealed containers. In these applications, it may not be necessary to actually measure the amount of radiation passing through the material, but only to detect changes in the level of radiation, in other words, for a "go, no-go" type of inspection.

When a highly absorbing material such as thick lead or concrete must be inspected for voids and when usual radiographic techniques are impractical, radiation gaging can be used effectively. Voids can be located in

these materials by noting increases in the readings of radiation detecting instruments.

Computed Tomography (CT). Cross-sectional images of an inspection object can be obtained by a series of radiation attenuation measurements all around the object. Typically, a fan beam of radiation about 1 or 2 mm in height is used along with a bank of detectors on the opposite side of the object. The attenuation data permit a computer reconstruction showing density differences in the cross section of the object (Fig. 9a).

X-ray computed tomography has many of the same benefits and limitations as film and real-time radiography. The primary difference is the nature of the radiological image. Radiography (Fig. 9b) compresses the structural information from a three-dimensional volume into a two-dimensional image. This is useful in that it allows a relatively large volume to be interrogated and represented in a single image. However, this compression limits the information and reduces the sensitivity to small variations. Radiographic images also can be difficult to interpret because of shadows from overlying and underlying structures superimposed on the features of interest. In contrast, the CT method provides sufficient information to localize a feature (Fig. 9a).

Some of the performance characteristics of radiography and computed tomography are compared in Table 5.

One of the limitations of CT inspection is that the CT image provides detailed information only over the limited volume of the cross-sectional slice. Full inspection of the entire volume of a component with computed tomography requires many slices, limiting the inspection throughput of the system. Therefore, CT equipment is often used in a digital radiography (DR) mode during production operations, with the CT imaging mode used for specific critical areas or to obtain more detailed information on indications found in the DR image. Digital radiography capabilities and throughput can be significant operational considerations for the overall system usage. Computed tomography systems generally provide a DR imaging mode, producing a two-dimensional radiographic image of the overall test piece.

Characteristics of X-Ray Film

Three general characteristics of film (speed, gradient, and graininess) are primarily responsible for the performance of the film during exposure and processing and for the quality of the resulting image. Film speed, gradient, and graininess are interrelated; that is, the faster the film, the larger the graininess and the lower the gradient, and vice versa. Film speed and gradient are derived from the characteristic curve for a film emulsion, which is a plot of film density versus the exposure required for producing that density in the processed film. Graininess is an inherent property of the

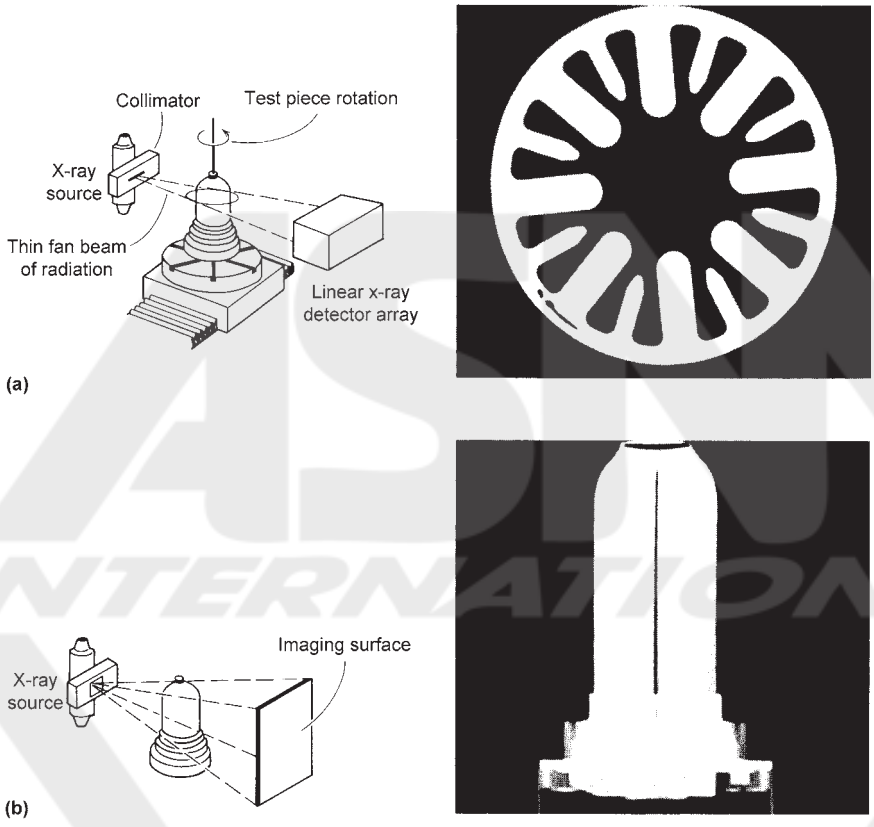


Fig. 9 Comparison of (a) computed tomography (CT) system and a CT image at the height of the flaw shows the flaw in more detail and in a form an inexperienced viewer can readily recognize; (b) radiography system and a high quality digital radiograph of a solid rocket motor igniter shows a serious flaw in a carefully oriented tangential shot. Source: Ref 1

emulsion, but can be influenced somewhat by the conditions of exposure and development.

The selection of radiographic film for a particular application is generally a compromise between the desired quality of the radiograph and the cost of exposure time. This compromise occurs because slower films generally provide a higher film gradient and a lower level of graininess and fog.

Film Types. The classification of radiographic film is complicated, as evidenced by changes in ASTM standard practice E94. The 1988 sequent editions of ASTM E94 references ASTM E746, which describes a standard test method for determining the relative image quality response of industrial radiographic film. Careful study of ASTM E746 is required to arrive at a conclusive classification index suitable for the given radiographic film requirements of a facility.

Earlier editions (1984 and prior) of ASTM E94 contained a table listing the characteristics of industrial films grouped into four types. The general characteristics of these four types are summarized in Table 6. This relatively simple classification method is referenced by many codes and specifications, which may state only that a type 1 or 2 film can be used for their specification requirements. However, because of this relatively arbitrary method of classification, many film manufacturers may be reluctant to assign type numbers to a given film. Moreover, the characteristics of radiographic films can vary within a type classification in Table 6 because of inherent variations among films produced by different manufacturers under different brand names and because of variations in film processing that affect both film speed and radiographic density. These variations make it essential that film processing be standardized and that characteristic curves for each brand of film be obtained from the film manufacturer for use in developing exposure charts.

Because the variables that govern the classification of film are no longer detailed in ASTM E94, it is largely the responsibility of film manufacturers to determine the particular type numbers associated with their brand names. Some manufacturers indicate the type number together with the brand name on the film package. If there is doubt regarding the type number of a given brand, it is advisable to consult the manufacturer. Most manufacturers offer a brand of film characterized as very low speed, ultrahigh gradient, and extremely fine grain.

Film selection for radiography is a compromise between the economics of exposure (film speed and latitude) and the quality desired in the radiograph. In general, fine grain, high gradient films produce the highest quality radiographs. However, because of the low speed typically associated with these films, high-intensity radiation or long exposure times are needed. Other factors affecting radiographic quality and film selection are the type and thickness of the test piece and the photon energy of the incident radiation.

Although the classification of film is more complex than the types given in Table 6, a general guide is that better radiographic quality will be promoted by the lowest type number in Table 6 that economic and technical

Table 6 General characteristics of the four types of radiographic film specified in the earlier (1984) edition of ASTM E94

	Film characteristic		
Film type	Speed	Gradient	Graininess
1	Low	Very high	Very fine
2	Medium	High	Fine
3	High	Medium	Coarse
4(a)	Very high(b) Medium(d)	Very high(b) Medium(d)	(c) Medium(d)

(a) Normally used with fluorescent screens. (b) When used with fluorescent screens. (c) Graininess is mainly a characteristic of the fluorescent screens. (d) When used for direct exposure or with lead screens. These groupings are given only for qualitative comparisons. For a more detailed discussion on film classification, see the section "Film Types" in this chapter. Source: Ref 1

considerations will allow. In this regard, Table 7 suggests a general comparison of film characteristics for achieving a reasonable level of radiographic quality for various metals and radiation source energies. However, it should be noted that the film types are only a qualitative ranking of the general film characteristics given in Table 6. Many radiographic films, particularly those designed for automatic processing, cannot be adequately classified according to the system in Table 6. This compounds the problem of selecting film for a particular application.

Exposure Factors

Exposure is the intensity of radiation multiplied by the time during which it acts; that is, the amount of energy that reaches a particular area of

Table 7 Guide to the selection of radiographic film for steel, aluminum, bronze, and magnesium in various thicknesses

Thickness		Type of film(a) for use with these x-ray tube voltages, or radioactive isotopes:										
mm	in.	50–80 kV	80–120 kV	120–250 kV	150–250 kV	Ir-192	250–400 kV	1 MeV	Co-60	2 MeV	Ra	6–31 MeV
Steel												
0–6	0–¼	3	3	2	1	...	...	...	...	...	...	...
6–13	¼–½	4	3	2	2	...	1	...	...	...	...	...
13–25	½–1	...	4	3	2	2	2	1	...	1	2	...
25–50	1–2	...	...	...	3	2	2	1	2	1	2	1
50–100	2–4	...	...	...	4	3	4	2	2	2	3	1
100–200	4–8	...	...	...	...	...	4	3	3	2	3	2
>200	>8	...	...	...	...	...	...	...	...	3	...	2
Aluminum												
0–6	0–¼	1	1	...	...	...	...	...	...	...	...	...
6–13	¼–½	2	1	1	1	...	...	...	...	...	...	...
13–25	½–1	2	1	1	1	...	1	...	...	...	...	...
25–50	1–2	3	2	2	1	1	1	...	...	...	...	...
50–100	2–4	4	3	2	2	1	2	...	...	...	...	...
100–200	4–8	...	4	3	3	2	3	...	...	...	...	...
>200	>8	...	...	...	...	4	...	...	...	...	...	...
Bronze												
0–6	0–¼	4	3	2	1	1	1	1	...	...	...	...
6–13	¼–½	...	3	2	2	2	1	1	...	1	...	...
13–25	½–1	...	4	4	3	2	2	1	2	1	2	...
25–50	1–2	...	...	4	4	3	3	1	2	1	2	1
50–100	2–4	...	...	...	...	3	4	2	3	2	3	1
100–200	4–8	...	...	...	...	...	...	3	3	2	...	2
>200	>8	...	...	...	...	...	...	...	...	3	...	2
Magnesium												
0–6	0–¼	1	1	...	...	...	...	...	...	...	...	...
6–13	¼–½	1	1	1	...	...	...	...	...	...	...	...
13–25	½–1	2	1	1	...	1	...	...	...	...	...	...
25–50	1–2	2	1	1	1	1	...	...	...	...	...	...
50–100	2–4	3	2	2	1	2	...	...	...	...	...	...
100–200	4–8	...	3	2	2	3	...	...	...	...	...	...
>200	>8	...	...	...	4	...	...	...	...	...	...	...

(a) These recommendations represent a usually acceptable level of radiographic quality and are based on the qualitative classification of films defined in Table 6. Optimum radiographic quality will be promoted by use of the lowest-number film type that economic and technical considerations will allow. The recommendations for type 4 film are based on the use of fluorescent screens. Source: Ref 1

the film and that is responsible for producing a particular density on the developed film. Density is the quantitative measure of blackening of a photographic emulsion. Density, measured directly with an instrument called a densitometer, is the logarithm of the ratio of the light intensity incident on the film to that transmitted by the film. Therefore, a film with a density of 1.0 will transmit only 10% of the light, a film with a density of 2.0 will transmit only 1/100 of the light, and so on.

There are two kinds of density: (a) the density associated with transparent base radiographic film, called transmission density, and (b) the density associated with opaque base imaging material such as radiographic paper, called reflection density.

The exposure time in film radiography depends mainly on film speed, the intensity of radiation at the film surface, the characteristics of any screens used, and the desired level of photographic density. In practice, the energy of the radiation is first chosen to be sufficiently penetrating for the type of material and thickness to be inspected. The film type and the desired photographic density are then selected according to requirements for contrast sensitivity. Once these factors are fixed, then the source strength, the source-to-film distance, and the characteristics of any screens used determine the exposure time.

With a given type of film and screen, the exposure time to produce the desired photographic density can be determined. Because intensity is inversely proportional to the square of distance from the source, the reciprocity law for equivalent exposures with an x-ray tube can be written as:

$$\frac{i_1 t_1}{L_1^2} = \frac{i_2 t_2}{L_2^2}$$

where i is the tube current, t is the exposure time, L is the source-to-film distance, and the subscripts refer to two different combinations that produce images with the desired photographic density. The parallel expression that applies to exposures made with a γ-ray source is:

$$\frac{a_1 t_1}{L_1^2} = \frac{a_2 t_2}{L_2^2}$$

where a is the source strength in gigabecquerel (Curies).

Contrast sensitivity refers to the ability of responding to and displaying small variations in subject contrast. Contrast sensitivity depends on the characteristics of the image detector and on the level of radiation being detected or on the amount of exposure for films. The relationship between the contrast sensitivity and the level of radiation intensity, or film exposure, can be illustrated by considering two extremes. At low levels of radiation intensity, the contrast sensitivity of the detectors is reduced by a smaller signal-to-noise ratio, while at high levels, the detectors become

saturated. Consequently, contrast sensitivity is a function of dynamic range (see below).

In film radiography, the contrast sensitivity is:

$$\text{Contrast sensitivity } \% = \frac{2.3\Delta D}{G_D} \times 100$$

where ΔD is the smallest change in photographic density that can be observed when the film is placed on an illuminated screen. The factor G_D is called the film gradient or film contrast. The film gradient is the inherent ability of a film to record a difference in the intensity of the transmitted radiation as a difference in photographic density. It depends on film type, development procedure, and film density. For all practical purposes, it is independent of the quality and distribution of the transmitted radiation.

Contrast sensitivity in real-time systems is determined by the number of bits (if the image is digitized) and the signal-to-noise ratio that is affected by the intensity of the radiation and the efficiency of the detector. The best contrast sensitivity of digitized images from fluorescent screens is about 8 bits (or 256 gray levels).

Another way of specifying the contrast sensitivity of fluorescent screens is with a gamma factor, which is defined by the fractional unit change in screen brightness, $\Delta B/B$, for a given fractional change in the radiation intensity, $\Delta I/I$. Most fluorescent screens have a gamma factor of about one, which is not a limiting factor. At low levels of intensity; however, the contrast is reduced because of quantum mottle (which is a form of screen unsharpness). Unsharpness may reduce the contrast depending on flaw size (Fig. 10).

Dynamic range, or latitude, describes the ability of the imaging system to produce a suitable signal over a range of radiation intensities. The dynamic range is given as the ratio of the largest signal that can be measured to the precision of the smallest signal that can be discriminated. A large dynamic range allows the system to maintain contrast sensitivity over a wide range of radiation intensities or test piece thicknesses. Film radiography has a dynamic range of up to 1000 to 1, while digital radiography with discrete detectors can achieve 100,000 to 1.

The latitude, or dynamic range, of film techniques is the range of test piece thickness that can be recorded with a single exposure. High gradient films generally have narrow latitude, that is, only a narrow range of test piece thickness can be imaged with optimum density for interpretation. If the test piece is of nonuniform thickness, more than one exposure may have to be made (using different x-ray spectra or different exposure times) for complete inspection of the piece. The number of exposures, as well as the exposure times, can often be reduced by using a faster film of lower gradient but wider latitude, although there is usually an accompanying reduction in ability to image small flaws.

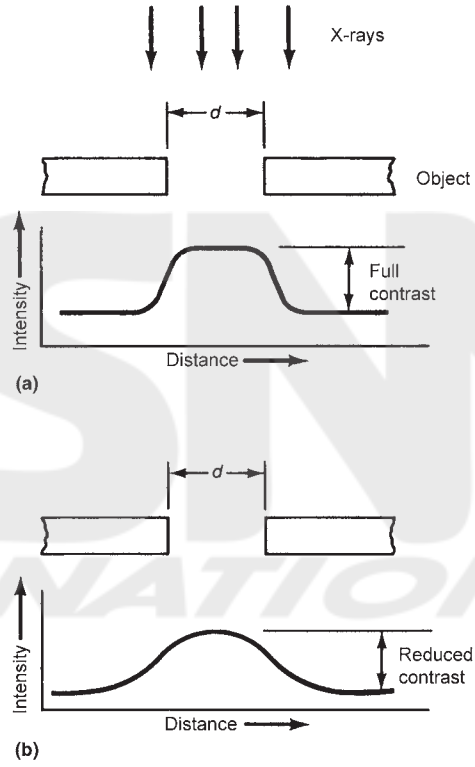


Fig. 10 Effect of geometric unsharpness on image contrast. (a) Flaw size *d* is larger than the unsharpness, then full contrast occurs. (b) Flaw size *d* is smaller than the unsharpness, then contrast is reduced. Source: Ref 1

Exposure Charts for X-Ray Radiography. Equipment manufacturers usually publish exposure charts for each type of x-ray generator that they manufacture. These published charts, however, are only approximations; each particular unit and each installation is unique. Radiographic density is affected by such factors as radiation spectrum, film processing, setup technique, amount and type of filtration, screens, and scattered radiation.

Although published exposure charts are acceptable guides for equipment selection, more accurate charts that are prepared under normal operating conditions are recommended for each x-ray machine. Simple steps for preparing accurate exposure charts are:

1. Make a series of radiographs of a calibrated multiple thickness step wedge, using several different values of exposure at each of several different tube voltage settings
2. Process the exposed films together under conditions identical to those that will be used for routine application
3. From the several densities corresponding to different thicknesses, determine which density (and thickness) corresponds exactly with the density desired for routine application. This step must be done with a

densitometer because no other method is accurate. If the desired density does not appear on the radiograph, the thickness corresponding to the desired density can be found by interpolation

4. Using the thickness determined in step 3 and the tube voltage (kilovoltage) and exposure (milliamp second or milliamp min) corresponding to that piece of film, plot the relation of thickness to exposure on semilogarithmic paper with exposure on the logarithmic scale
5. Draw lines of constant tube voltage through the corresponding points on the graph

Spectral Sensitivity. The shape of the characteristic curve of a given x-ray film is for all practical purposes unaffected by the wavelength distribution in the x-ray or gamma ray beam used for the exposure. However, the sensitivity of the film in terms of roentgens required to produce a given density is strongly affected by radiation energy (beam spectrum of a given kilovoltage or given gamma ray source).

Figure 11 shows the exposure required for producing a density of 1.0 on type 4 radiographic film developed in an x-ray developer (made from powder) for five minutes at 20 °C (68 °F). The exposures were made directly, without screens. The spectral sensitivity curves for all x-ray films have approximately the same general features as the curves shown in Fig. 11.

The classification of radiographic film is complicated; however, a relatively simple classification has been adopted by ASTM. According to the classification in ASTM E94, radiographic films are grouped into four types. The general characteristics of these four types are summarized in Table 6. The relative image quality of x-ray film can also be determined in a quantitative manner using a multihole test piece as described in ASTM E746.

Screens are often used with x-ray films during exposure. Metal screens, typically used at x-ray energies of over 150 kV, intensify the image by emission of photoelectrons and help reduce the effects of scatter by atten-

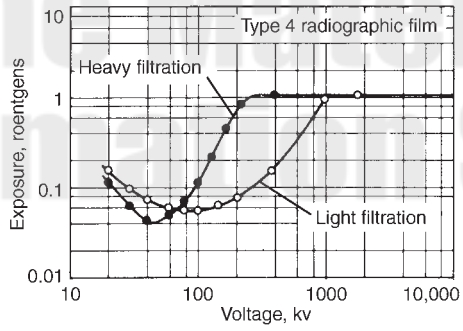


Fig. 11 Spectral sensitivity curves for a type 4 radiographic film, showing exposure required to produce a density of 1.0. Source: Ref 1

uating the lower-energy scattered radiation. Lead is typically used. Fluorescent screens are used in some situations to help reduce exposure times.

Image Quality. The quality level of an industrial radiograph is governed by the radiographic sensitivity exhibited on the radiograph itself. Radiographic sensitivity is determined through the use of penetrameters or image-quality indicators (IQI).

Penetrameters, or IQIs, are of a known size and shape, and have the same attenuation characteristics as the material in the test piece. They are placed on the test piece or on a block of identical material during setup and are radiographed at the same time as the test piece. Penetrameters are preferably located in regions of maximum test piece thickness and greatest test piece-to-film distance, and near the outer edge of the central beam of radiation. The degree to which features of the penetrameter are visible in the developed image is a measure of the quality of that image. The image of the penetrameter that appears on the finished radiograph is evaluated during interpretation to ensure that the desired sensitivity, definition, and contrast have been achieved in the developed image.

Penetrameters of different designs have been developed by various standards making organizations. Common types are plaques containing holes and a second type containing a series of wires. Plaque type penetrameters consist of strips of material of uniform thickness with holes drilled through them specified by ASTM E142. Wire type penetrameters are widely used in Europe (German standard DIN 54109). They are also used in the United States and are described in ASTM standard E747. The sensitivity of a wire type penetrameter is expressed in terms of wire diameter divided by object thickness.

Neutron Radiography

Neutron radiography is a form of nondestructive inspection that uses a specific type of particulate radiation, called neutrons, to form a radiographic image of a test piece. The geometric principles of shadow formation, variation of attenuation with test piece thickness, and many other factors that govern the exposure and processing of a neutron radiograph are similar to those for radiography using x-rays or gamma rays.

The section deals mainly with the characteristics that differentiate neutron radiography from x-ray or gamma ray radiography. The application of neutron radiography is described in terms of its advantages for improved contrast on low atomic number materials, discrimination between isotopes, or inspection of radioactive specimens.

Neutrons are subatomic particles that are characterized by relatively large mass and a neutral electric charge. The attenuation of neutrons differs from the attenuation of x-rays in that the processes of attenuation are nuclear rather than processes that depend on interaction with electron shells surrounding the nucleus.

Neutrons are produced by nuclear reactors, accelerators, or certain radioactive isotopes, all of which emit neutrons of relatively high energy (fast neutrons). Because most neutron radiography is performed with neutrons of lower energy (thermal neutrons), the sources are usually surrounded by a *moderator*, which is a material that reduces the kinetic energy of the neutrons by scattering.

Neutron radiography differs from conventional radiography in that the attenuation of neutrons as they pass through the test piece is more related to the specific isotope present than to density or atomic number. X-rays are attenuated more by elements of high atomic number than by elements of low atomic number, and this effect varies relatively smoothly with atomic number. Also, x-rays are generally attenuated more by materials of high density than they are by materials of low density. For thermal neutrons, the attenuation tends to decrease with increasing atomic number, although the trend is by no means a smooth relationship. In addition to the high attenuation of several light elements (hydrogen, lithium, and boron), certain medium to heavy elements (especially cadmium, samarium, europium, gadolinium, and dysprosium) and certain specific isotopes have exceptionally high capabilities for absorbing thermal neutrons. This means that neutron radiography is capable of detecting these highly attenuating elements or isotopes when present in a structure of lower absorption capability.

Using neutrons, it is possible to radiographically detect certain isotopes, for instance, certain isotopes of hydrogen, cadmium, or uranium. Some neutron image detection methods are insensitive to gamma rays or x-rays and can be used to inspect radioactive materials such as reactor fuel elements. The high attenuation of hydrogen, in particular, opens many application possibilities, including inspection of assemblies for detection of adhesives, explosives, lubricants, water, hydrides, corrosion, plastics, or rubber.

Neutron Sources

The excellent discrimination capabilities of neutrons generally refer to neutrons of low energy, that is, thermal neutrons. Characteristics of neutron radiography corresponding to various ranges of neutron energy are summarized in Table 8. Although any of these energy ranges can be used for radiography, this section emphasizes the thermal neutron range, which is the most widely used for inspection.

In thermal-neutron radiography, an object or test piece is placed in a thermal-neutron beam in front of an image detector. The neutron beam may be obtained from a nuclear reactor, a radioactive source, or an accelerator. Several characteristics of these sources are summarized in Table 9. For thermal neutron radiography, fast neutrons emitted by these sources must first be moderated and then collimated. The radiographic intensities

Table 8 Characteristics of neutron radiography at various neutron energy ranges

Type of neutrons	Energy range	Characteristics
Cold	Below 0.01 eV	High absorption cross sections decrease transparency of most materials, but also increase efficiency of detection. An advantage is reduced scatter at energies below the Bragg cutoff, where neutrons can no longer undergo Bragg reflection.
Thermal	0.01 – 0.3 eV	Good discrimination between materials and ready availability of sources.
Epithermal	0.3 eV – 10 keV	Excellent discrimination for particular materials by working at energy of resonance. Greater transmission and less scatter in samples containing materials such as hydrogen and enriched reactor fuels.
Fast	10 keV – 20 MeV	Good point sources are available. At low energy end of spectrum, fast neutron radiography may be able to perform many inspections done with thermal neutrons, but with a panoramic technique. Good penetration capability because of low absorption cross sections in all materials. Poor material discrimination

Source: Ref 1

Table 9 Several characteristics of thermal neutron sources

Type of source	Typical radiographic intensity(a)	Resolution	Exposure time	Characteristics
Radioisotope	10^1–10^4	Poor to medium	Long	Stable operation, low to medium investment cost, possibly portable
Accelerator	10^3–10^6	Medium	Average	On-off operation, medium cost, possibly transportable
Subcritical assembly	10^4–10^6	Good	Average	Stable operation, medium to high investment cost, movement difficult
Nuclear reactor	10^5–10^8	Excellent	Short	Medium to high investment cost, movement difficult

(a) Neutrons per cm^2 per second. Source: Ref 1

listed in Table 9 typically do not exceed 10^{-5} times the total fast neutron yield of the source. Part of this loss is incurred in moderating the neutrons, and the remainder in bringing a collimated beam out of a large volume moderator.

Collimation is necessary for thermal neutron radiography because there are no useful point sources of low-energy neutrons. Good collimation in thermal neutron radiography is comparable to small focal spot size in conventional radiography; the images of thick objects will be sharper with good collimation. Conversely, it should be noted that available neutron intensity decreases with increasing collimation.

Neutron Detection Methods

Detection methods for neutron radiography generally make use of photographic or x-ray films. In the direct-exposure method, film is exposed directly to the neutron beam, with a conversion screen or intensifying screen providing secondary radiation that actually exposes the film. Alternatively, film can be used to record an autoradiographic image from a radioactive image carrying screen in a technique called the transfer method.

Direct Exposure Method. Conversion screens of thin gadolinium foil or a scintillator have been most widely used in the direct exposure method. When bombarded with a beam of neutrons, some of the gadolinium atoms absorb neutrons and promptly emit gamma rays and internal conversion electrons. Scintillators are fluorescent screens, often made of zinc sulfide crystals that also contain a specific isotope such as $_3Li^6$ or $_5B^{10}$. Gadolinium oxysulfide, a scintillator originally developed for x-ray radiography, has been widely used for neutron radiography.

Scintillators provide useful images with total exposures as low as 5×10^5 neutrons per cm^2. The high speed and favorable relative response make scintillators attractive for use with nonreactor neutron sources. Gadolinium screens provide greater uniformity and image sharpness (high-contrast resolution of 10 μm has been reported), but an exposure about 30 or more times that of a scintillator is required, even with fast films.

Transfer Method. In the transfer method, a thin sheet of metal, typically of indium or dysprosium, is exposed to the neutron beam transmitted through the specimen. Neutron capture induces radioactivity−indium having a half-life of 54 minutes and dysprosium a half-life of 2.35 hours. The *radiograph* to be interpreted is made by placing the radioactive transfer screen in contact with a sheet of film. Beta-particles and gamma rays from the transfer screen expose the film.

The transfer method is especially valuable for inspection of a radioactive specimen. Although radiation emitted by the specimen (especially gamma rays) causes heavy film fogging during x-ray radiography or direct exposure neutron radiography, the same radiation will not induce radioactivity in a transfer screen. Thus, a clear image of the specimen can be obtained even when there is a high level of background radiation.

In comparing the two primary detection methods, the direct exposure method offers high speed, indefinite image integration time and the best spatial resolution. The transfer method offers insensitivity to gamma rays emitted by the specimen and greater contrast because of lower amounts of scattered and secondary radiation.

Real time imaging, in which light from a scintillator is observed by a television camera, also can be used for neutron radiography. Because of low brightness, most real-time neutron radiographic images are enhanced by an image intensifier tube, which may be separate or integral with a television camera. This method can be used for applications such as the study of fluid flow in a closed system such as a heat pipe or engine or the study of metal flow in a mold during casting.

Applications

Various applications that are discussed in ASTM STP 586 emphasize the value of neutron radiography for inspection of ordnance, explosive,

aerospace, and nuclear components. The presence, absence, or correct placement of explosives, adhesives, O-rings, plastic components, and similar materials can be verified. The presence of fluids or corrosion can be detected. Nuclear fuel and control materials can be inspected to determine distribution of isotopes and to detect foreign or imperfect material. Hydride deposition in metals and diffusion of boron in heat-treated, boron-fiber composites can be observed.

The characteristics of neutron radiography complement those of conventional x-radiography; one radiation provides a capability lacking or difficult for the other.

ACKNOWLEDGMENT

This chapter was adapted from Radiography in *Metals Handbook Desk Edition*, Second Edition, 1998.

REFERENCES

1. Radiography, *Metals Handbook Desk Edition*, 2nd ed., ASM International, 1998, p 1292–1302

SELECTED REFERENCES

- L. Cartz, *Nondestructive Testing*, ASM International, 1995
- D.E. Bray and R.K. Stanley, *Nondestructive Evaluation: A Tool in Design, Manufacturing, and Service*, Taylor & Francis, 1996
- *Nondestructive Testing and Quality Control*, Vol 17, *ASM Handbook*, ASM International, 1989
- P.J. Shull, *Nondestructive Evaluation: Theory, Techniques, and Applications*, Marcel Dekker, 2001

CHAPTER 11

Ultrasonic Inspection

ULTRASONIC INSPECTION is a nondestructive method in which beams of high frequency acoustic energy are introduced into a material to detect surface and subsurface flaws, to measure the thickness of the material, and to measure the distance to a flaw. An ultrasonic beam travels through a material until it strikes an interface or discontinuity such as a flaw. Interfaces and flaws interrupt the beam and reflect a portion of the incident acoustic energy. The amount of energy reflected is a function of (a) the nature and orientation of the interface or flaw and (b) the acoustic impedance of such a reflector. Energy reflected from various interfaces and flaws can be used to define the presence and locations of flaws; the thickness of the material; and, the depth of a flaw beneath a surface.

Most ultrasonic inspections are performed using a frequency between 1 and 25 MHz. Short shock bursts of ultrasonic energy are aimed into the material from the ultrasonic search unit of the ultrasonic flaw detector instrument. The electrical pulse from the flaw detector is converted into ultrasonic energy by a piezoelectric transducer element in the search unit. The beam pattern from the search unit is determined by the operating frequency and size of the transducer element. Ultrasonic energy travels through the material at a specific velocity that is dependent on the physical properties of the material and on the mode of propagation of the ultrasonic wave. The amount of energy reflected from or transmitted through an interface, other type of discontinuity, or reflector depends on the properties of the reflector. These phenomena provide the basis for establishing two of the most common measurement parameters used in ultrasonic inspection: the amplitude of the energy reflected from an interface or flaw and the time required from pulse initiation for the ultrasonic beam to reach the interface or flaw.

Ultrasonic Flaw Detectors

Although the electronic equipment used for ultrasonic inspection can vary greatly in detail among equipment manufacturers, all general purpose units consist of a power supply, a pulser circuit, a search unit, a receiver-amplifier circuit, an oscilloscope, and an electronic clock. Many systems also include electronic equipment for signal conditioning, gating, automatic interpretation, and integration with a mechanical or electronic scanning system. Also, advances in microprocessor technology have extended the data acquisition and signal processing capabilities of ultrasonic inspection systems. In most instances, the entire electronic assembly, including the controls and display, is contained in a single instrument. A typical ultrasonic flaw detector is shown in Fig. 1. The major controls included are:

- Frequency selector to select the operating or test frequency
- Pulse tuning control to fine adjust the test frequency
- Pulse repetition rate control, which determines the number of times per second that an ultrasonic pulse is initiated from the transducer (typically 100 to 2000 pulses per second)
- Test type or mode selection switch to adjust instrument to pulse echo or pitch catch operation
- Sensitivity controls to adjust sensitivity or gain of the receiver-amplifier
- Sweep selector and delay to adjust time base and that portion of the inspection zone that is to be displayed
- Gate position control to isolate the portion of the inspection zone that will be used for additional processing

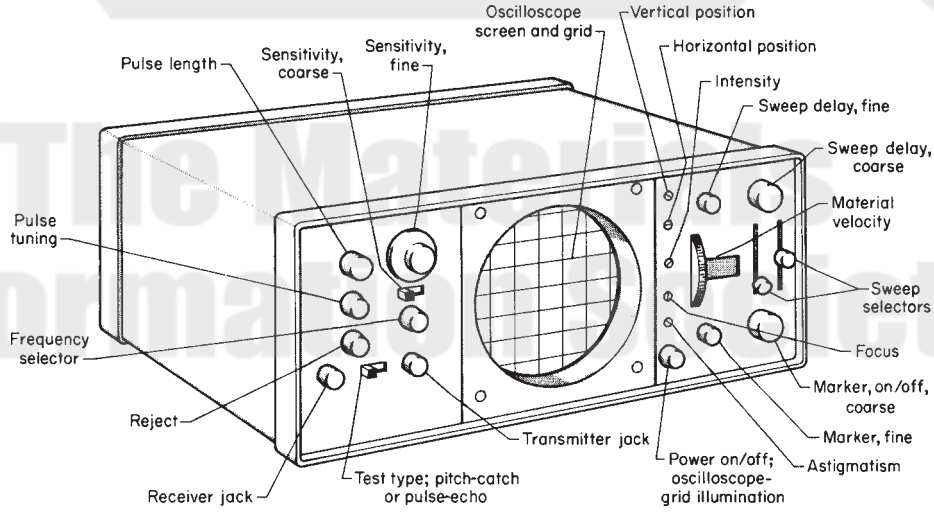


Fig. 1 Typical pulse echo instrument. Source: Ref 1

- Oscilloscope, which provides the visual display of the time and amplitude parameters used to interpret the data from the ultrasonic inspection

Ultrasonic Transducers and Search Units

Generation and detection of ultrasonic waves for inspection is accomplished by means of a transducer element. The transducer element is contained within a device most often referred to as a search unit (or probe). The active element in a search unit is a piezoelectric crystal. Piezoelectricity is pressure induced electricity, a property characteristic of certain naturally occurring crystalline compounds and some man made materials. An electrical charge is developed by the crystal when pressure is applied to it. Conversely, when an electrical field is applied, the crystal mechanically deforms (changes shape). Piezoelectric crystals have various deformation modes with thickness expansion being the principal mode used in transducers for ultrasonic inspection.

The most common types of piezoelectric materials used for ultrasonic search units are quartz; lithium sulfate; and polarized ceramics such as barium titanate, lead zirconate titanate, and lead metaniobate. The characteristics and applications of these materials are summarized in Table 1.

Search units come in a variety of types and shapes. Variations in search-unit construction include transducer element material, transducer element thickness, surface area, and shape; and type of backing material and degree of loading. Four basic types of search units are straight beam contact, angle beam contact, dual element contact, and immersion, both flat and focused. Their primary areas of application are listed in Table 2. Sectional views of these search units, together with a special type (delay-tip, contact-type search unit), are shown in Fig. 2.

The selection of a transducer depends very much on the properties of the test specimen, particularly its sound attenuation. Ultrasonics of high frequency produce good resolution, which is the ability to separate echoes

Table 1 Characteristics and applications of piezoelectric transducer elements

	Characteristics of piezoelectric elements									
	Efficiency							Suitability of element in:		
			Coupling					Contact inspection		
Piezoelectric element	Transmit	Receive	To water	To metal	Tolerance to elevated temperature	Damping ability	Undesired modes (inherent noise)	Straight-beam	Angle-beam	Immersion inspection
Quartz	P	G	G	F	G	F	G	G	F	G
Lithium sulfate	F	E	E	P	P	E	E	P	F	E
Barium titanate	G	P	G	G	P	P	P	G	G	F
Lead zirconate titanate	E	F	F	E	E	F	P	E	E	F
Lead metaniobate	G	F	G	E	E	E	G	E	E	G

E, excellent; G, good; F, fair; P, poor. Source: Ref 1

Table 2 Primary applications of ultrasonic search units

Unit type	Manufacturing induced flaws	Service induced flaws
Straight beam, contact	Billets: inclusions, stringers, pipe Forgings: inclusions, cracks, segregations, seams, flakes, pipe Rolled products: laminations, inclusions, tears, seams, cracks Castings: slag, porosity, cold shuts, tears, shrinkage cracks, inclusions	Fatigue cracks, corrosion, erosion, stress-corrosion cracks
Angle beam, contact	Forgings: cracks, seams, laps Rolled products: tears, seams, cracks, cupping Welds: slag inclusions, porosity, incomplete fusion, incomplete penetration, dropthrough, suckback, cracks in filler metal and base metal Tubing and pipe: circumferential and longitudinal cracks	Fatigue cracks, stress-corrosion cracks
Dual element, contact	Plate and sheet: thickness measurement, lamination detection Tubing and pipe: measurement of wall thickness	Wall thinning, corrosion, erosion, stress-corrosion cracks
Immersion	Billets: inclusions, stringers, pipe Forgings: inclusions, cracks, segregations, seams, flakes, pipe Rolled products: laminations, inclusions, tears, seams, cracks Welds: inclusions, porosity, incomplete fusion, incomplete penetration, dropthrough, cracks, base-metal laminations Adhesive-bonded, soldered, or brazed products: lack of bond Composites: voids, resin rich, resin poor, lack of filaments Tubing and pipe: circumferential and longitudinal cracks	Corrosion, fatigue cracks

Source: Ref 1

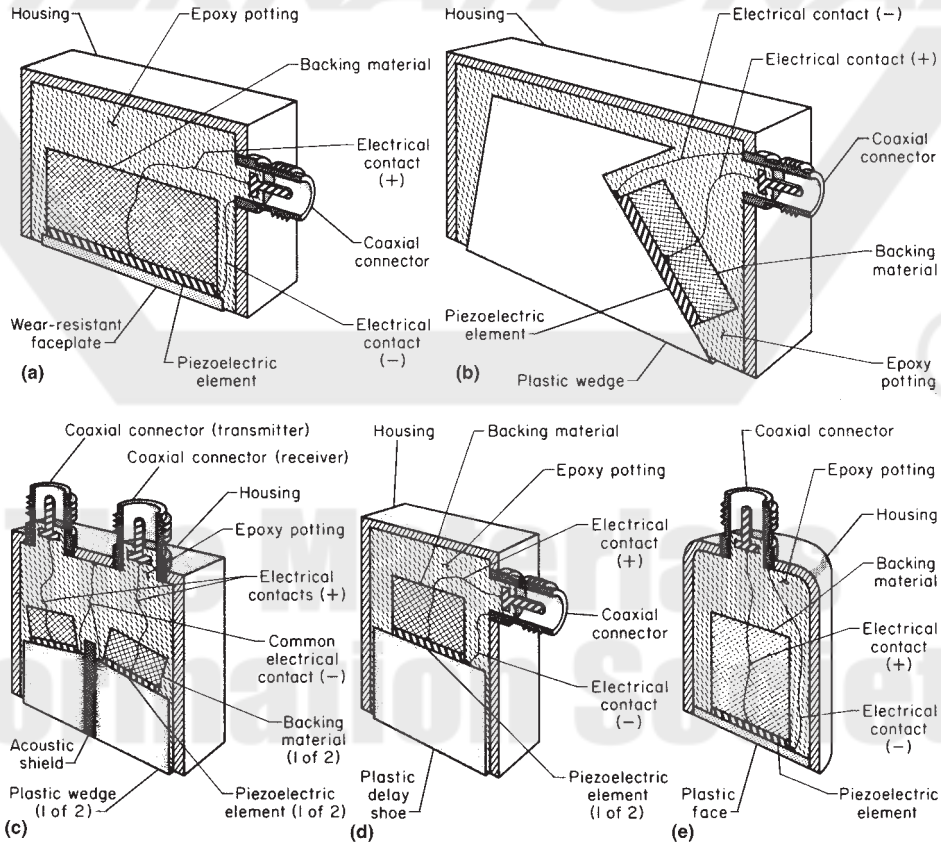


Fig. 2 Sectional views of five types of search units used in ultrasonic inspection: (a) straight-beam (longitudinal-wave) contact, (b) angle-beam (shear-wave) contact, (c) dual-element contact, (d) delay-tip (stand-off) contact, and (e) immersion. Source: Ref 1

from closely spaced defects. Ultrasonics of low frequency penetrate deeper into materials because attenuation is generally lower. However, backscattering "noise" from grain boundaries is usually more important than attenuation, although the net result is the same because the signal-to-noise ratio also usually decreases with frequency, as shown in Fig. 3. This means that the two requirements of high penetration and high resolution are mutually exclusive. For example, a specimen having high attenuation, such as steel, should be examined by a low frequency beam of about 0.5 MHz and a large transducer diameter about 50 mm (2 in.), which provides a high penetration, but a relatively low lateral resolution of about 6 mm (1/4 in.). Improved resolution can be obtained by using shear waves, because these have shorter wavelengths than compression waves of the same frequency in the solid. The velocity of a longitudinal wave in a solid is greater than that of the shear wave of the same frequency. Large diameter transducers are chosen to produce a narrow focused beam, which enhances the lateral resolution.

Couplants

Air is a poor transmitter of sound waves at megahertz frequencies. Also, because the acoustic impedance mismatch between air and most solids is significant, even a very thin layer of air severely retards the transmission of sound waves from the transducer to the test piece. Therefore, it is necessary to use a couplant to eliminate air between the transducer and the test piece for satisfactory contact inspection.

Couplants normally used for contact inspection include water, oils, glycerin, petroleum greases, silicone grease, cellulose gum, and various commercial pastelike substances. Certain soft rubbers that transmit sound

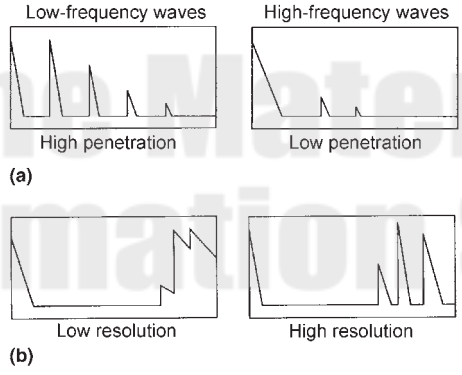


Fig. 3 Oscilloscope displays using ultrasonic transducers of (a) high and low penetration (ability to detect defects at distances within the solid), and (b) high and low resolution (ability to separate echoes from closely spaced defects). Source: Ref 1

waves can be used where adequate coupling can be achieved by applying pressure to the search unit.

Factors that should be considered in selecting a couplant are:

- Surface finish of test piece
- Temperature of test surface
- Possibility of chemical reactions between test surface and couplant
- Cleaning requirements (some couplants are difficult to remove)

Water is a suitable couplant for use on a relatively smooth surface, but a wetting agent should be added. The addition of glycerin sometimes is necessary to increase viscosity.

Heavy oil or grease should be used on hot and vertical surfaces and on rough surfaces where irregularities need to be filled.

Cellulose gum is especially useful on rough surfaces when good coupling is needed to minimize background noise and yield an adequate signal-to-noise ratio.

Basic Inspection Methods

Ultrasound can be used to measure material thickness by (a) determining resonant frequencies of a test piece and (b) measuring time required for an ultrasonic wave packet (pulse) to traverse the test piece. The former uses reflected ultrasound to create standing waves in the test piece; the frequencies at which standing waves occur are used to compute thickness. In the latter method, the time it takes for a pulse of ultrasonic energy to be transmitted through the test piece is measured; this time period can be 100 nanoseconds or less. Thickness is calculated as the product of the measured time of flight and the known acoustic wave velocity.

Ultrasound can be used to detect flaws by measuring (a) the amplitude of the acoustic pressure wave and time of flight of reflected acoustic waves and (b) the amplitude of the acoustic pressure wave of either transmitted or reflected acoustic waves. The pulse echo technique is the most widely used ultrasonic technique. Flaws are detected and their sizes estimated by comparing the amplitude of a reflected echo from an interface (either within the test piece or at the back surface) with the amplitude of an echo reflected from a reference interface of known size or from the back surface of a test piece that has no flaws. The echo from the back surface (back reflection) serves as a reference point for time-of-flight measurements that enable measuring the depth of some internal flaws. It is necessary that an internal flaw reflect at least part of the sound energy onto the receiving transducer to measure depth. However, echoes from flaws are not essential to their detection. Just because the amplitude of an echo from back reflection of a test piece is lower than that of an echo from an identical flaw-free workpiece implies that the test piece contains one or more flaws.

Detection of the presence of flaws by sound attenuation is used in both transmission and pulse echo techniques. The inability to detect flaw depth is the main disadvantage of attenuation techniques.

Pulse Echo Method

In pulse echo inspection, short bursts of ultrasonic energy (pulses, or wave packets) are introduced into a test piece at regular time intervals. If the pulses encounter a reflecting surface, some or all of the energy is reflected. The proportion of energy that is reflected is highly dependent on the size of the reflecting surface in relation to the size of the incident ultrasonic beam. The direction of the reflected beam or echo depends on the orientation of the reflecting surface with respect to the incident beam. Reflected energy is monitored; both the amount of energy reflected in a specific direction and the time delay between transmission of the initial pulse and receipt of the echo are measured.

Principles of Operation. Most pulse echo systems consist of (a) an electronic clock; (b) an electronic signal generator, or pulser; (c) a sending transducer; (d) a receiving transducer; (e) an echo signal amplifier; and, (f) a display device. In the most widely used version of pulse echo systems, a single transducer acts alternatively as a sending and receiving transducer. The clock and signal generator usually are combined in a signal electronic unit. Frequently, circuits that amplify and demodulate echo signals from the transducer are housed in the same unit.

In a pulse echo system with a single search unit, the electronic clock triggers the signal generator at regular intervals, which imposes a short burst of high frequency alternating voltage on the transducer element. Simultaneously, the clock activates a time measuring circuit connected to the display device. The operator preselects a constant interval between pulses by means of a *pulse repetition rate* control on the instrument; pulses usually are repeated 100 to 2000 times per second. The operator also can preselect the signal generator or pulser output frequency. For best results, frequency (and sometimes the pulse-repetition rate) should be tuned to achieve the maximum response of the transducer (resonance in the vibrating element) and maximum signal-to-noise ratio (lowest amount of electronic noise) in the electronic equipment. The transducer converts the pulse of alternating voltage into a pulse of mechanical vibration having essentially the same frequency as the imposed alternating voltage. The mechanical vibration (ultrasound) is introduced into a test piece through a couplant and travels by wave motion through the test piece at the speed of sound. When the pulse of ultrasound encounters a reflecting surface that is perpendicular to the direction of travel, ultrasonic energy is reflected and returns to the transducer. The returning pulse travels along the same path and at the same speed as the initial pulse, but in the opposite direction.

Data Presentation. Information from pulse echo inspection can be displayed in one of three forms: (a) A-scan, which is a quantitative display of echo amplitude and time-of-flight data obtained at a single point on the surface of the test piece; (b) B-scan, which is a quantitative cross-sectional display of time-of-flight data obtained along a plane perpendicular to the surface of the test piece; or (c) C-scan, which is a semiquantitative display of echo amplitude obtained over an area of the surface of the test piece. The A-scan display, which is the most widely used form, can be analyzed in terms of the type, size, and location (chiefly depth) of internal flaws.

A-scan display basically is a plot of amplitude versus time, in which a horizontal baseline on an oscilloscope screen indicates elapsed time and vertical deflections called *indications* or *signals* represent echos. A typical A-scan setup that illustrates the essential elements in a basic system for pulse echo inspection is shown in Fig. 4. These elements are:

- Power supply, which can run on 110-volt alternating current or on batteries
- Electronic clock or timing circuit, to trigger pulser and display circuits
- Pulser circuit or rate generator, to control frequency, amplitude, and pulse-repetition rate of the voltage pulses that excite the search unit
- Receiver-amplifier circuit to convert output signals from the search unit into a form suitable for oscilloscope display
- Sweep circuit to control (a) time delay between search-unit excitation

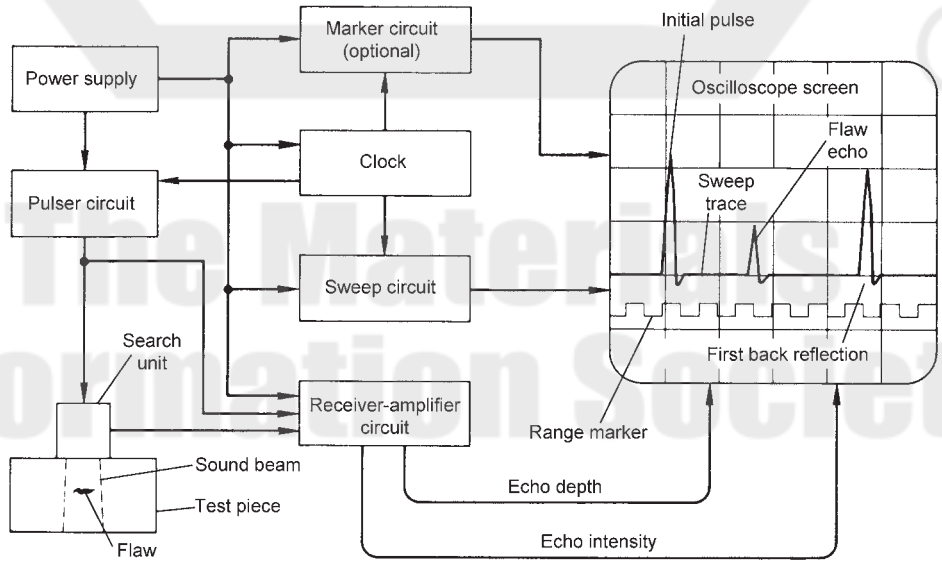


Fig. 4 Typical A-scan setup including video mode display for a basic pulse echo ultrasonic inspection system. Source: Ref 1

and start of oscilloscope trace and (b) rate at which oscilloscope trace travels horizontally across the screen
- Marker circuit (optional) to produce a secondary trace, on or below the main trace, usually in the form of a square wave, which is used for precise depth measurements
- Oscilloscope screen, including separate controls for trace brightness, trace focus, and illuminated measuring grid
- Flaw gate (not shown) to isolate the echo of interest for further processing

The search unit and the coaxial cable connecting the unit to the instrument, although not strictly part of the electronic circuitry, must be matched to the electronics.

B-scan display is a plot of time versus distance, in which one orthogonal axis on the display corresponds to elapsed time, while the other axis represents the position of the search unit along a line on the surface of the test piece relative to the position of the search unit at the start of the inspection. Echo amplitude is not measured directly as in A-scan inspection, but often it is indicated semiquantitatively by the relative brightness of echo indications on an oscilloscope screen.

A typical B-scan system is shown in Fig. 5. System functions are identical to those of the A-scan system except for the following differences:

- The display is generated on an oscilloscope screen consisting of a long-persistence phosphor; that is, a phosphor that continues to fluoresce long after the means of excitation ceases to fall on the fluoresc-

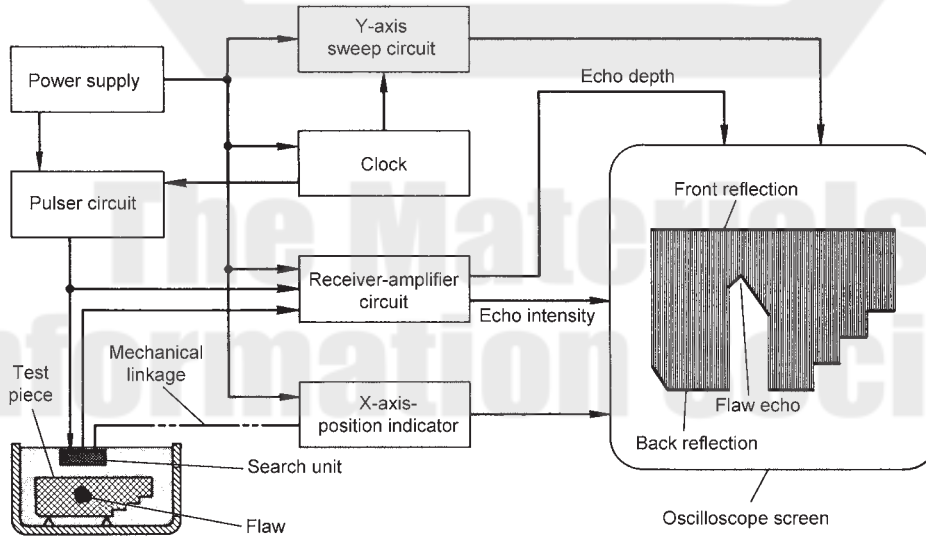


Fig. 5 Typical B-scan setup including video mode display for a basic pulse echo ultrasonic inspection system. Source: Ref 1

ing area of the screen. This allows the imaginary cross section to be viewed as a whole without having to resort to permanent imaging methods such as photographs. (Photographic equipment, facsimile recorders, or *x-y* plotters can be used to record B-scan data for later reference.)

- Oscilloscope input for one axis of the display is provided by an electromechanical device, which generates an electrical voltage proportional to the position of the search unit relative to a reference point on the surface of the test piece. Most B-scans are generated by scanning the search unit in a straight line across the surface of the test piece at a uniform rate. One axis of the display (usually the horizontal axis) represents the distance traveled along this line.
- Echoes are indicated by bright spots on the screen rather than by deflections of the time trace. The position of a bright spot along the axis orthogonal to the search-unit position axis (usually measured top to bottom on the screen) indicates the depth of the echo within the test piece.
- The echo-intensity signal from the receiver-amplifier is connected to the trace-brightness control on the oscilloscope to ensure that echoes are recorded as bright spots. In some systems, the brightnesses corresponding to different values of echo amplitude has sufficient contrast to permit semiquantitative appraisal of echo amplitude, which is related to flaw size and shape.

The oscilloscope screen in Fig. 5 illustrates the type of video-mode display that is generated by B-scan equipment. The internal flaw in the test piece shown at left in Fig. 5 is shown only as a profile view of its top reflecting surface. Portions of the test piece that are behind this large reflecting surface are in shadow.

C-scan display records echoes from internal portions of test pieces as a function of the position of each reflecting interface within an area. Flaws are shown on a readout superimposed on a plan view of the test piece, and both flaw size (flaw area) and position within the plan view are recorded. Flaw depth typically is not recorded, although it can be measured semiquantitatively by restricting the range of depths within the test piece that is covered in a given scan.

In a basic C-scan system, shown schematically in Fig. 6, the search unit is moved over the surface of the test piece in a search pattern. The search pattern can take many forms, such as a series of closely spaced parallel lines, a fine zigzag pattern, and a spiral pattern (polar scan). Mechanical linkage connects the search unit to *x*-axis and *y*-axis position indicators, which in turn, feed position data to the *x-y* plotter or facsimile device. Echo-recording systems vary; some produce a shaded-line scan with echo amplitude recorded as a variation in line shading, while others indicate

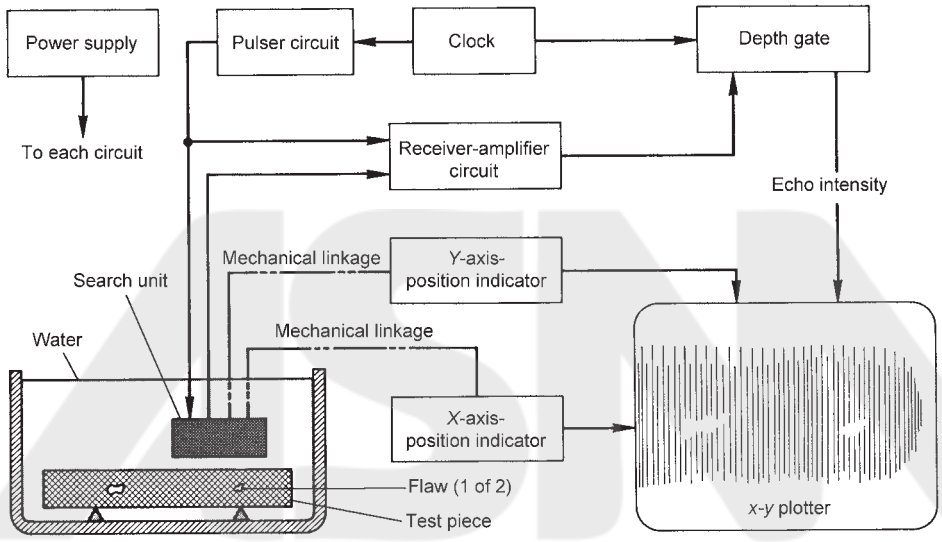


Fig. 6 Typical C-scan setup including display for a basic pulse echo ultrasonic inspection system. Source: Ref 1

flaws by an absence of shading, so each flaw appears as a blank space on the display.

Interpretation of Pulse Echo Data

Interpretation of pulse echo data is relatively straightforward for B-scan and C-scan presentations. The B-scan always records the front reflection, while internal echoes and/or loss of back reflection are interpreted as flaw indications. Flaw depth is measured as the distance from the front reflection to a flaw echo, the latter representing the front surface of the flaw.

In contrast to normal B-scan and C-scan displays, A-scan displays can be complex. It is necessary to disregard electronic noise, spurious echoes, and extra echoes resulting from mode conversion of the initial pulse to focus attention on any flaw echoes that might be present.

Basic A-scan displays are of the type shown in Fig. 7 for immersion inspection of a plate containing a flaw. The test material is 25 mm (1 in.) thick aluminum alloy 1100 plate containing a purely reflecting planar flaw 11.25 mm (0.44 in.) deep. The flaw is 45% of plate thickness, exactly parallel to the plate surfaces, and has an area equal to one-third the cross section of the ultrasonic beam. Testing is by straight beam immersion in a water filled tank. There are no attenuation losses within the test plate, only transmission losses across front and back surfaces.

The normal display (Fig. 7c) represents only a portion of the complete video mode A-scan display (Fig. 7b). The normal display is obtained by adjusting horizontal position and horizontal sweep controls to display

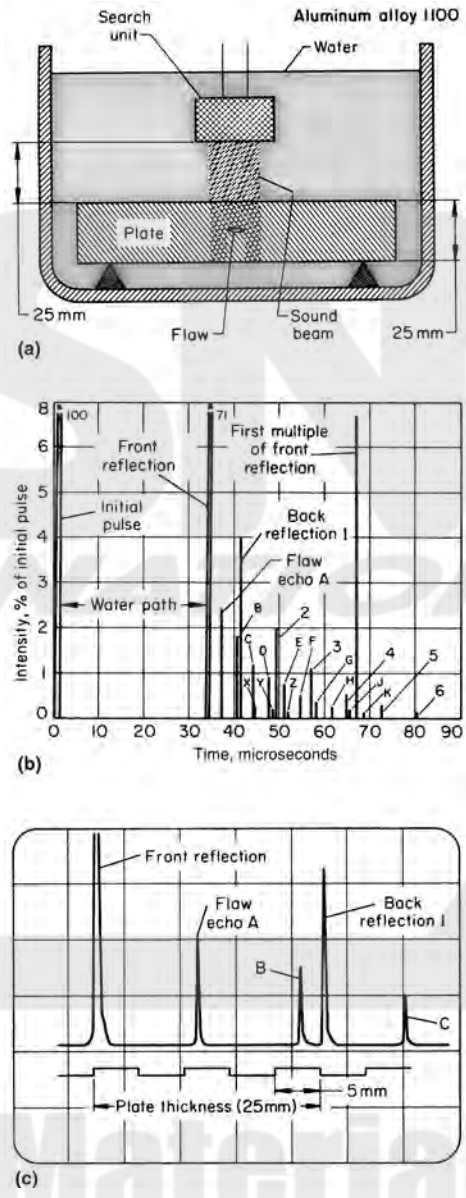


Fig. 7 Schematic representation of straight-beam immersion inspection of a 25 mm thick aluminum alloy 1100 plate containing a planar discontinuity, showing (a) inspection setup, (b) complete video mode A-scan display, and (c) normal oscilloscope display. Source: Ref 1

only the portion of the trace corresponding to the transit time (time of flight) required for a single pulse of ultrasound to traverse the test piece from front surface to back surface and return. Also, receiver amplifier gain is adjusted until the height of the first back reflection equals some arbitrary distance on the screen, usually a convenient number of grid lines.

Most flaws are not exactly parallel to the surface of the test piece, not truly planar (they have rough, curved interfaces), not ideal reflectors, and are of unknown size. These factors together with bulk material sound attenuating characteristics affect echo signal size and shape.

Angle Beam Technique. Most angle beam testing is accomplished using shear waves, although refracted longitudinal waves and surface waves can be used in some applications. In contrast to straight beam testing, only flaw indications appear on the display in an angle-beam test. Only rarely will a back surface be oriented properly to give a back reflection indication. In most instances, ultrasonic beams are reflected from the back surface at an angle away from the search unit. The reflected pulses are capable of detecting discontinuities and are used extensively in angle-beam testing of welds, pipe and tubing, and sheet and plate.

The time base (horizontal sweep) on the oscilloscope must be carefully calibrated, because in angle-beam testing there is no back reflection echo to provide a reference to estimate flaw depth. Usually, an extended time base is used so flaws are located with one or two skip distances from the search unit. The definition of skip distance is shown in Fig. 8.

Figure 8(a) shows how a shear wave from an angle-beam transducer progresses through a flat test piece by reflecting from the surfaces at points called *nodes*. The linear distance between two successive nodes on the same surface is called the *skip distance*, and is important in defining the path over which the transducer should be moved for reliable and efficient scanning. The skip distance can easily be measured by using a separate receiving transducer to detect the nodes, or by using an angle-beam test block, or it can be calculated. The region over which the transducer should be moved to scan the test piece can be determined once the skip distance is known.

Moving the search unit back and forth between one-half skip distance and one skip distance from an area of interest can be used not only to de-

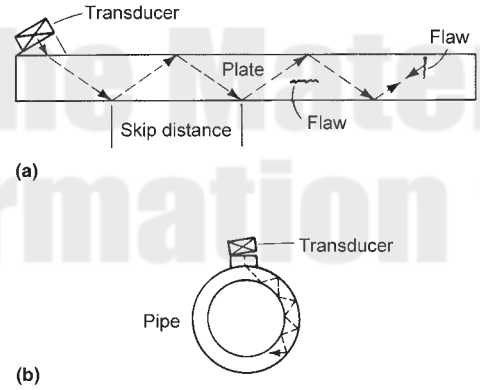


Fig. 8 Angle-beam testing using a contact transducer on a (a) plate and (b) pipe. Source: Ref 1

fine the location, depth, and size of a flaw, but also to initially detect flaws. This back-and-forth movement as a way of scanning a weld for flaws is illustrated in Fig. 9.

Sometimes, moving the search unit in an arc about the position of a suspected flaw or swiveling the search unit about a fixed position can be equally useful (Fig. 10a). As shown in Fig. 10(b), traversing the search unit in an arc about the location of a gas hole produces little or no change in the echo. The indication on the oscilloscope screen remains constant in both amplitude and position on the trace as the search unit is moved. Conversely, if the search unit was swiveled on the same spot, the indication would abruptly disappear after the search unit was swiveled only a few degrees.

Transmission Methods

Regardless of whether transmission ultrasonic testing is done using direct beams or reflected beams, flaws are detected by comparing the amount

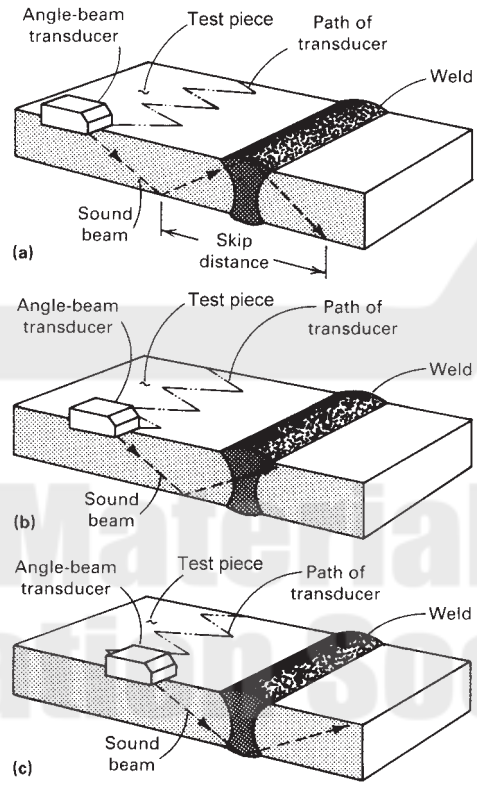


Fig. 9 Three positions of a contact type of transducer along the zigzag scanning path used during manual angle-beam ultrasonic inspection of welded joints. The movement of the sound beam path across the weld is shown on a section taken along the centerline of the transducer as it is moved from the far left position in the (a) scanning path, (b) through an intermediate position, (c) to the far right position. Source: Ref 1

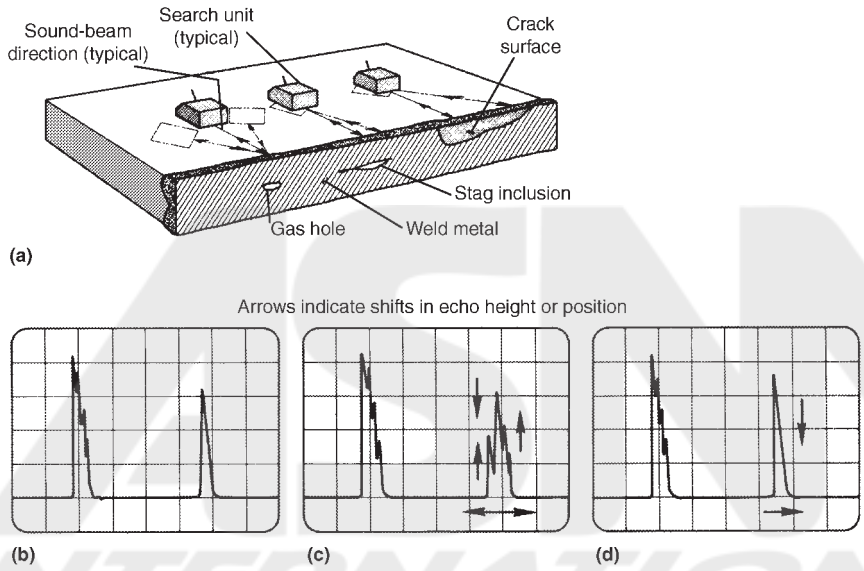


Fig. 10 Angle-beam inspection of a weldment, showing effect of search unit movements on oscilloscope screen display patterns from three different types of flaws in welds. (a) Positions of search units on the test piece. (b) Display pattern obtained from a gas hole as the result of traversing the search unit in an arc about the location of the flaw. (c) Display pattern obtained from a slag inclusion as the result of swiveling the search unit on a fixed point. (d) Display pattern obtained from a crack, using the same swiveling search-unit movement as in (c). Source: Ref 1

of ultrasound transmitted through the test piece with the amount transmitted through a reference standard made of the same material. Transmission testing requires two search units, one to transmit the ultrasonic waves and one to receive them.

The main application of transmission methods is the inspection of plate for cracks or laminations that have relatively large dimensions compared with the size of the search units. Immersion techniques and water-column (bubbler or squirter) techniques are most effective because they provide efficient and relatively uniform coupling between the search units and the test piece.

Display of transmission-test data can be oscilloscope traces, strip-chart recordings, and meter readings. Oscilloscopes are used to record data mainly when using pulsed sound beams; strip charts and meters are more appropriate for continuous beams. With all three types of display, alarms or automatic sorting devices can be used to give audible warning or to shunt defective workpieces out of the normal flow of production.

Pitch-catch testing can be done with either direct beams (through transmission testing) or reflected beams. In both instances, pulses of ultrasonic energy pass through the material, and pulse intensities are measured at the point of emergence. An oscilloscope display is triggered simultaneously with the initial pulse, and the transmitted pulse indication appears on the screen to the right of the initial pulse indication in a manner similar

to the back reflection indication in pulse echo testing. A major advantage of pitch-catch testing is that disturbances and spurious indications can be separated from the transmitted pulse by their corresponding transit times. Only the amplitude of the transmitted pulse is monitored; all other sound waves reaching the receiver are ignored. An electronic gate can be set to operate an alarm or a sorting device when the monitored amplitude of the ultrasonic wave drops below a preset value.

When reflected pulses are used, the technique is almost identical to the loss of back reflection technique, which often is used in ordinary pulse echo testing.

General Characteristics of Ultrasonic Waves

In contrast to electromagnetic waves, such as light and x-rays, ultrasonic waves are mechanical waves consisting of oscillations, or vibrations, of the atomic or molecular particles of a substance about the equilibrium position of those particles. Ultrasonic waves can propagate in elastic media, which can be solid or liquid. Ultrasonic waves in the megahertz region are severely attenuated in air and cannot propagate in a vacuum. An ultrasonic beam is similar to a light beam. Both obey general wave equations and each travels at a characteristic velocity that depends on the properties of that medium. Ultrasonic beams, like light beams, are reflected from surfaces and are refracted when they cross boundaries between two media that have different acoustic velocities. Depending on the mode of particle motion, ultrasonic waves are classified as longitudinal waves, vertically and horizontally polarized shear and transverse waves, surface waves, Lamb waves, etc. Four wave modes are described in the following paragraphs.

Longitudinal waves, sometimes called compression waves, are most widely used in the inspection of metals. They travel through metal as a series of alternate compressions and rarefactions, in which the particles transmitting the wave vibrate back and forth in the direction of travel of the waves.

Longitudinal ultrasonic waves and the corresponding particle oscillation and resultant rare-faction and compression are represented schematically in Fig. 11(a). A plot of amplitude of particle displacement versus distance of wave travel, together with the resultant rarefaction trough and compression crest, is shown in Fig. 11(b). The distance from one crest to the next (which equals the distance for one complete cycle of rarefaction and compression) is the wavelength (λ). The vertical axis in Fig. 11(b) could represent pressure instead of particle displacement. The horizontal axis could represent time instead of travel distance because the speed of sound is constant in a given material, and this relation is used in the measurements made in ultrasonic inspection.

Longitudinal ultrasonic waves are readily propagated in liquids and elastic solids. The mean free paths of the molecules of liquids are so short that longitudinal waves can be propagated simply by the elastic collision of one molecule with the next. The velocity of longitudinal ultrasonic waves is about 6000 m/s (19,700 ft/s) in steel and about 1500 m/s (4900 ft/s) in water.

Transverse waves (shear waves) also are used extensively in ultrasonic inspection of metals. Transverse waves are visualized readily in terms of vibrations of a rope that is shaken rhythmically, in which each particle vibrates up and down in a plane perpendicular to the direction of propagation. A transverse wave is represented schematically in Fig. 12, which shows particle oscillation, wave front, direction of wave travel, and the wavelength (λ) corresponding to one cycle.

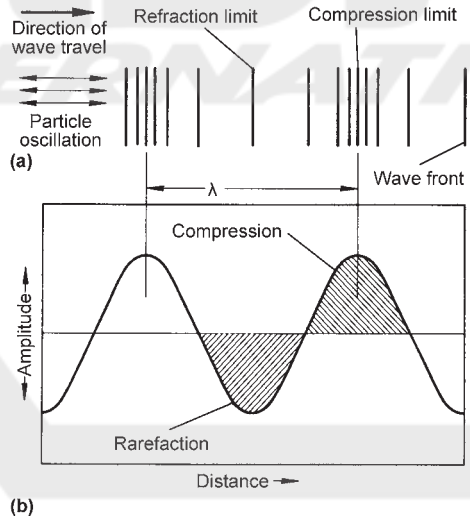


Fig. 11 Schematic representation of longitudinal ultrasonic waves. (a) Particle oscillation and resultant rarefaction and compression. (b) Amplitude of particle displacement versus distance of wave travel. The wavelength (λ) is the distance corresponding to one complete cycle. Source: Ref 1

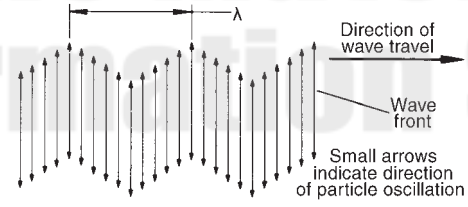


Fig. 12 Schematic representation of transverse or shear waves. The wavelength (λ) is the distance corresponding to one complete cycle. Source: Ref 1

Air and water do not support transverse waves. In gases, the forces of attraction between molecules are so small that shear waves cannot be transmitted. The same is true of a liquid, unless it is particularly viscous or is present as a very thin layer.

Surface waves (Rayleigh waves) are another type of ultrasonic waves used in the inspection of metals. These waves travel along the flat and curved surfaces of relatively thick solid parts. For propagation of waves of this type, the waves must be traveling along an interface bounded on one side by the strong elastic forces of a solid, and on the other side by the practically negligible elastic forces between gas molecules. Surface waves, therefore, are essentially nonexistent in a solid immersed in a liquid, unless the liquid covers the solid surface only as a very thin film.

Surface waves are subject to less attenuation in a given material than are longitudinal and transverse waves. They have a velocity approximately 90% of the transverse-wave velocity in the same material. The region within which these waves propagate with effective energy is not much thicker than about one wavelength beneath the surface of the metal. At this depth, wave energy is about 4% of the wave energy at the surface, and the amplitude of oscillation decreases sharply to a negligible value at greater depths.

In Rayleigh waves, particle oscillation generally follows an elliptical orbit, as shown schematically in Fig. 13. The major axis of the ellipse is perpendicular to the surface along which the waves are traveling. The minor axis is parallel to the direction of propagation. Rayleigh waves can exist in complex forms, which are variations of the simplified wave form illustrated in Fig. 13.

Lamb waves, also known as *plate waves*, are propagated in a mode in which the ultrasonic beam is contained within two parallel boundary surfaces (such as a plate or the wall of a tube). A Lamb wave consists of a complex vibration that occurs throughout the thickness of the material. The propagation characteristics of Lamb waves depend on the density, elastic properties, and structure of the metal, and are influenced by material thickness.

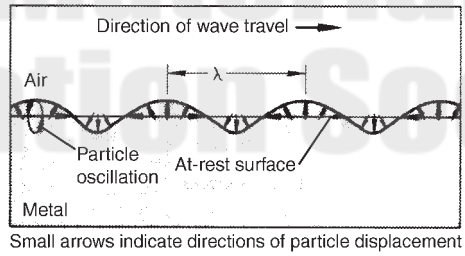


Fig. 13 Diagram of surface (Rayleigh) waves propagating at the surface of a metal along a metal-air interface. The wavelength (λ) is the distance corresponding to one complete cycle. Source: Ref 1

Two basic forms of Lamb waves are (a) symmetrical or dilatational, and (b) asymmetrical or bending. The form is determined by whether the particle motion is symmetrical or asymmetrical with respect to the neutral axis of the test piece. Each form is further subdivided into several modes having different velocities, which can be controlled by the angle at which the waves enter the test piece. Theoretically, there are an infinite number of specific velocities at which Lamb waves can travel in a given material. Within a given plate, the specific velocities of Lamb waves are complex functions of plate thickness and cyclic frequency.

In symmetrical Lamb waves, there is a compressional (longitudinal) particle displacement along the neutral axis of the plate and an elliptical particle displacement on each surface (see Fig. 14a). In asymmetrical Lamb waves, there is a shear (transverse) particle displacement along the neutral axis of the plate and an elliptical particle displacement on each surface (see Fig. 14b). The ratio of the major to minor axes of the ellipse is a function of the material in which the wave is being propagated.

Factors Influencing Ultrasonic Inspection

Both the characteristics of ultrasonic waves used and the part being inspected must be considered in ultrasonic inspection. Equipment type and

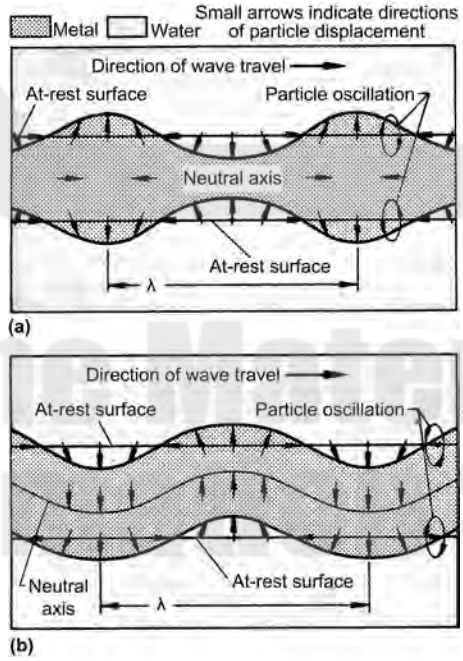


Fig. 14 Diagram of the basic patterns of (a) symmetrical (dilatational) and (b) asymmetrical (bending) Lamb waves. The wavelength (λ) is the distance corresponding to one complete cycle. Source: Ref 1

capability are influenced by these variables; often, different types of equipment must be selected to accomplish different inspection objectives.

Selection of inspection frequency is a compromise between the ability of the ultrasonic beam to penetrate the material and the time or depth resolution desired. A high frequency generally provides high resolution and high definition, while a lower frequency might be required to achieve the desired penetration.

Sensitivity, or the ability of an ultrasonic-inspection system to detect a very small discontinuity, generally is increased by using relatively high frequencies (short wavelengths). Frequency ranges commonly used in nondestructive testing (NDT) are listed in Table 3.

Acoustic Impedance. When ultrasonic waves traveling through one medium impinge on the boundary of a second medium; a portion of the incident acoustic energy is reflected back from the boundary while the remaining energy is transmitted into the second medium. The characteristic that determines the amount of reflection is the acoustic impedance of the two materials on either side of the boundary. If the impedances of the two materials are equal, there is no reflection; if the impedances differ greatly (between a metal and air, for example), there is virtually complete reflection.

This characteristic is used in ultrasonic inspection of metals to calculate the amounts of energy reflected and transmitted at impedance discontinuities, and to aid in the selection of suitable materials for effective transfer of acoustic energy between components in ultrasonic-inspection systems.

The acoustic impedance for a longitudinal wave (Z_l), in grams per square centimeter-second, is defined as the product of material density (ρ), in grams per cubic centimeter, and longitudinal-wave velocity (V_l), in centimeters per second:

$$Z_l = \rho V_l$$

The acoustic properties of several metals and nonmetals are listed in Table 4. The acoustic properties of metals and alloys are influenced by variations in structure and metallurgical condition. Therefore, for a given test piece, the properties may differ somewhat from the values shown in Table 4.

Table 3 Common ultrasonic testing frequency ranges and applications

Frequency range	Applications
200 kHz–1 MHz	Coarse-grain castings: gray iron, nodular iron, copper, and stainless steels
400 kHz–5 MHz	Fine-grain castings: steel, aluminum, brass
200 kHz–2.25 MHz	Plastics and plastic like materials
1–5 MHz	Rolled products: metallic sheet, plate, bars, and billets
2.25–10 MHz	Drawn and extruded products: bars, tubes, and shapes
1–10 MHz	Forgings
2.25–10 MHz	Glass and ceramics
1–2.25 MHz	Welds
1–10 MHz	Fatigue cracks

Source: Ref 1

Table 4 Acoustic properties of several metals and nonmetals

Material	Density (ρ), g/cm³	Sonic velocities, 10^5 cm/s			Acoustic impedance(d), (Z_t) 10^6 g/cm² · s
		V_l(a)	V_l(b)	V_s(c)	
Ferrous metals					
Carbon steel, annealed	7.85	5.94	3.24	3.0	4.66
Alloy steel					
Annealed	7.86	5.95	3.26	3.0	4.68
Hardened	7.8	5.90	3.23	…	4.6
Cast iron	6.95–7.35	3.5–5.6	2.2–3.2	…	2.5–4.0
52100 steel					
Annealed	7.83	5.99	3.27	…	4.69
Hardened	7.8	5.89	3.20	…	4.6
D6 tool steel					
Annealed	7.7	6.14	3.31	…	4.7
Hardened	7.7	6.01	3.22	…	4.6
Stainless steels					
Type 302	7.9	5.66	3.12	3.12	4.47
Type 304L	7.9	5.64	3.07	…	4.46
Type 347	7.91	5.74	3.10	2.8	4.54
Type 410	7.67	5.39	2.99	2.16	4.13
Type 430	7.7	6.01	3.36	…	4.63
Nonferrous metals					
Aluminum 1100-O	2.71	6.35	3.10	2.90	1.72
Aluminum alloy 2117-T4	2.80	6.25	3.10	2.79	1.75
Beryllium	1.85	12.80	8.71	7.87	2.37
Copper C11000	8.9	4.70	2.26	1.93	4.18
Copper alloys					
C26000 (cartridge brass, 70%)	8.53	3.83	2.05	1.86	3.27
C46400 to C46700 (naval brass)	8.41	4.43	2.12	1.95	3.73
C51000 (phosphor bronze, 5% A)	8.86	3.53	2.23	2.01	3.12
C75200 (nickel silver 65–18)	8.75	4.62	2.32	1.69	4.04
Lead					
Pure	11.34	2.16	0.70	0.64	2.45
Hard (94Pb–6Sb)	10.88	2.16	0.81	0.73	2.35
Magnesium alloy M1A	1.76	5.74	3.10	2.87	1.01
Mercury, liquid	13.55	1.45	…	…	1.95
Molybdenum	10.2	6.25	3.35	3.11	6.38
Nickel					
Pure	8.8	5.63	2.96	2.64	4.95
Inconel	8.5	5.82	3.02	2.79	4.95
Inconel X-750	8.3	5.94	3.12	…	4.93
Monel	8.83	5.35	2.72	2.46	4.72
Titanium, commercially pure	4.5	6.10	3.12	2.79	2.75
Tungsten	19.25	5.18	2.87	2.65	9.98
Nonmetals					
Air(e)	0.00129	0.331	…	…	0.00004
Ethylene glycol	1.11	1.66	…	…	0.18
Glass					
Plate	2.5	5.77	3.43	3.14	1.44
Pyrex	2.23	5.57	3.44	3.13	1.24
Glycerin	1.26	1.92	…	…	0.24
Oil					
Machine (SAE 20)	0.87	1.74	…	…	0.150
Transformer	0.92	1.38	…	…	0.127
Paraffin wax	0.9	2.2	…	…	0.2
Plastics					
Methylmethacrylate (Lucite, Plexiglas)	1.18	2.67	1.12	1.13	0.32
Polyamide (nylon)	1.0–1.2	1.8–2.2	…	…	0.18–0.27
Polytetrafluoroethylene (Teflon)	2.2	1.35	…	…	0.30
Quartz, natural	2.65	5.73	…	…	1.52
Rubber, vulcanized	1.1–1.6	2.3	…	…	0.25–0.37
Tungsten carbide	10–15	6.66	3.98	…	6.7–9.9
Water					
Liquid(f)	1.0	1.49	…	…	0.149
Ice(g)	0.9	3.98	1.99	…	0.36

(a) Longitudinal (compression) waves. (b) Transverse (shear) waves. (c) Surface waves. (d) For longitudinal waves $Z_t = \rho V_l$. (e) At standard temperature and pressure. (f) At 4 °C (39 °F). (g) At 0 °C (32 °F). Source: Ref 1

Angle of Incidence. Only when an ultrasonic wave is incident at right angles on an interface between two materials (normal incidence, or angle of incidence = 0°) do transmission and reflection occur at the interface without any change in beam direction. At any other angle of incidence, the phenomena of mode conversion (a change in the nature of the wave motion) and refraction (a change in direction of wave propagation) must be considered. These phenomena can affect the entire beam or only a portion of the beam. The sum total of the changes that occur at the interface depends on the angle of incidence and the velocity of the ultrasonic waves leaving the point of impingement on the interface. All possible ultrasonic waves leaving this point are shown for an incident longitudinal ultrasonic wave in Fig. 15. Not all the waves shown in Fig. 15 will be produced in any specific instance of oblique impingement of an ultrasonic wave on an interface between two materials. The waves that propagate in a given instance depend on the angle of incidence of the initial beam, the velocities of the wave forms in both materials, and the ability of a wave form to exist in a given material.

Critical Angles. If the angle of incidence (α_1 in Fig. 15) is small, sound waves propagating in a given medium undergo mode conversion at a boundary, resulting in simultaneous propagation of longitudinal and transverse (shear) waves in a second medium. If the angle is increased, the direction of the refracted longitudinal wave will approach the plane of the boundary ($\alpha_2 \rightarrow 90°$). At some specific value of α_1, α_2 will exactly equal 90°, and the refracted longitudinal wave will disappear, leaving only a refracted (mode-converted) shear wave to propagate in the second medium. This value of α_1 is known as the *first critical angle*. If α_1 is increased beyond the first critical angle, the direction of the refracted shear wave will approach the plane of the boundary ($\beta_2 \rightarrow 90°$). At a second specific value

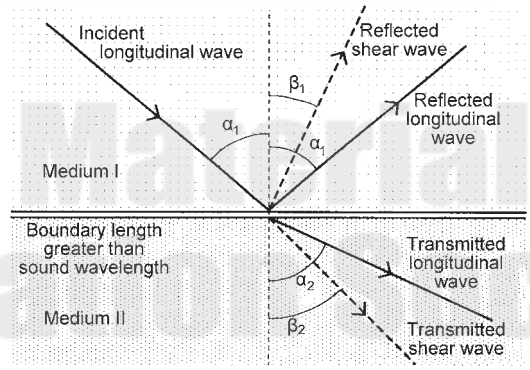


Fig. 15 Wave mode conversion at a boundary. There is an angle of incidence α_1 of the incoming longitudinal wave, such that the angle of the transmitted longitudinal wave α_2 becomes 90°. At angles of incidence greater than α_1, the longitudinal wave of velocity does not penetrate into medium II, and only the shear wave is transmitted. This is used to separate longitudinal and shear waves to have only a single wave velocity traveling in medium II. Source: Ref 1

of α_1, β_2 will exactly equal 90° and the refracted transverse wave will disappear. This second value of α_1 is called the *second critical angle*.

In ordinary angle beam inspection, it usually is desirable to have only a shear wave propagating in the test material. Because longitudinal waves and shear waves propagate at different speeds, echo signals are received at different times, depending on which type of wave produces the echo. When both types are present in the test material, confusing echo patterns can be displayed on the readout device, which can lead to an erroneous interpretation. Frequently, it can be useful to produce shear waves in a material at an angle of 45° to the surface. In most materials, incident angles for mode conversion to a 45° shear wave lie between the first and second critical angles. Typical values of α_1 for all three of these (first critical angle, second critical angle, and incident angle for mode conversion to 45° shear waves) are listed in Table 5 for various metals.

Absorption of ultrasonic energy occurs mainly by conversion of mechanical energy into heat. Elastic motion within a substance as a sound wave propagates through it alternately heats the substance during compression and cools it during rarefaction. Because heat flows so much more slowly than an ultrasonic wave, thermal losses are incurred, which progressively reduces energy in the propagating wave. A related thermal loss occurs in polycrystalline materials: a thermoelastic loss arises from heat flow away from grains that have received more compression or expansion in the course of wave motion than did adjacent grains. For most polycrystalline materials this effect is most pronounced at the low end of the ultrasonic-frequency spectrum.

Table 5 Critical angles for immersion and contact testing, and incident angle for 45° shear-wave transmission, in various metals

	First critical angle, degrees(a), for:		Second critical angle, degrees(a), for:		45° shear-wave incident angle, degrees(a), for:	
Metal	Immersion testing(b)	Contact testing(c)	Immersion testing(b)	Contact testing(c)	Immersion testing(b)	Contact testing(c)
Steel	14.5	26.5	27.5	55	19	35.5
Cast iron	15–25	28–50	. . .	. . .	. . .	. . .
Type 302 stainless steel	15	28	29	59	19.5	37
Type 410 stainless steel	11.5	21	30	63	20.5	39
Aluminum alloy 2117-T4	13.5	25	29	59.5	20	37.5
Beryllium	6.5	12	10	18	7	12.5
Copper alloy C26000 (cartridge brass, 70%)	23	44	46.5	. . .	31	67
Inconel	11	20	30	62	20.5	38.5
Magnesium alloy M1A	15	27.5	29	59.5	20	37.5
Monel	16.5	30	33	79	23	44
Titanium	14	26	29	59	20	37

(a) Measured from a direction normal to surface of test material. (b) In water at 4 °C (39 °F). (c) Using angle block (wedge) made of acrylic plastic. Source: Ref 1

Scattering of an ultrasonic wave occurs because most materials are not truly homogeneous. Crystal discontinuities such as grain boundaries, twin boundaries, and minute nonmetallic inclusions deflect small amounts of ultrasonic energy out of the main ultrasonic beam. Also, especially in mixed microstructures and anisotropic materials, mode conversion at crystallite boundaries occurs because of slight differences in acoustic velocity and acoustic impedance across the boundaries.

Scattering is highly dependent on the relation of crystallite size (mainly grain size) to ultrasonic wavelength. When grain size is less than 0.01 times the wavelength, scatter is negligible. Scattering effects vary approximately with the third power of grain size, and when the grain size is 0.1 times the wavelength or larger, excessive scattering may make it impossible to do valid ultrasonic inspections.

Diffraction. A sound beam propagating in a homogeneous medium is coherent; that is, all particles that lie along any given plane parallel to the wave front vibrate in identical patterns. When a wave front passes the edge of a reflecting surface, the front bends around the edge in a manner similar to that in which light bends around the edge of an opaque object. When the reflector is very small compared with the sound beam, as is usual for a pore or an inclusion, wave bending (forward scattering) around the edges of the reflector produces an interference pattern in a zone immediately behind the reflector because of phase differences among different portions of the forward-scattered beam. The interference pattern consists of alternate regions of maximum and minimum intensity that correspond to regions where interfering scattered waves are in phase and out of phase, respectively.

Diffraction phenomena must be taken into account during development of ultrasonic-inspection procedures.

Near-Field and Far-Field Effects. The face of the transducer element vibrates in a complex manner, which can most easily be described as a mosaic of tiny, individual crystals, each vibrating in the same direction but slightly out of phase with its neighbors. Each element in the mosaic functions as a point (Huygens) source, and radiates a spherical wave outward from the plane of the transducer face.

Along the central axis of the composite ultrasonic beam, the series of acoustic-pressure maximums and minimums become broader and more widely spaced as the distance from the transducer face, d, increases. Where d becomes equal to N (length of the near field), the acoustic pressure reaches a final maximum and decreases approximately exponentially with increasing distance, as shown in Fig. 16.

Beam Spreading. In the far field of an ultrasonic beam, the wave front expands with increasing distance from a transducer. The angle of divergence from the central axis of the beam from a circular transducer is determined from ultrasonic wavelength and transducer size.

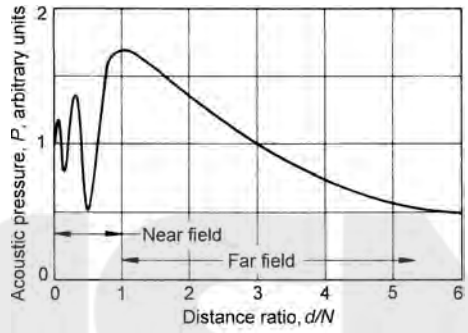


Fig. 16 Variation of acoustic pressure with distance ratio for a circular search unit. Distance ratio is distance from crystal face *d* divided by length of near field *N*. Source: Ref 1

Advantages, Disadvantages, and Applications

Advantages. The principal advantages of ultrasonic inspection compared with other methods of nondestructive inspection of metal parts are:

- Superior penetrating power, which permits detection of flaws deep in the part. Ultrasonic inspection is done routinely to depths of several feet in many types of parts and to depths of about 6 m (20 ft) in axial inspection of parts such as long steel shafts and rotor forgings
- High sensitivity, permitting detection of extremely small flaws
- Greater accuracy than other nondestructive methods in determining the positions of internal flaws, estimating their sizes, and characterizing them in terms of nature, orientation, and shape
- Only one surface need be accessible
- Operation is electronic, which provides almost instantaneous indications of flaws. This makes the method suitable for immediate interpretation, automation, rapid scanning, in-line production monitoring, and process control. With most systems, a permanent record of inspection results can be made
- Volumetric scanning ability, permitting inspection of a volume of metal extending from the front surface to the back surface of a part
- Ultrasonic inspection presents no radiation hazard to operations or nearby personnel, and has no effect on equipment and materials in the vicinity
- Portability

Disadvantages of ultrasonic inspection are:

- Manual operation requires careful attention by experienced technicians
- Technical knowledge is required to develop inspection procedures

- Parts that are rough, irregular in shape, very small and thin, and not homogeneous are difficult to inspect
- Discontinuities that are present in a shallow layer immediately beneath the surface might not be detectable
- Couplants are needed to provide effective transfer of the ultrasonic beam between search units and parts being inspected
- Reference standards are required, both to calibrate equipment and to characterize flaws

Applications. Some of the major types of components that are ultrasonically inspected for the presence of flaws are:

- Mill components: rolls, shafts, drives, and press columns
- Power equipment: turbine forgings, generator rotors, pressure piping, weldments, pressure vessels, nuclear fuel elements, and other reactor components
- Jet-engine parts: turbine and compressor forgings, and gear blanks
- Aircraft components: forging stock, frame sections, and honeycomb sandwich assemblies
- Machinery materials: die blocks, tool steels, and drill pipe
- Railroad parts: axles, wheels, and bolted and welded rail
- Automotive parts: forgings, ductile castings, brazed and/or welded components

Ultrasonic inspection is an effective and inexpensive method for volumetric examination of structures and components of both regular and complex shapes.

ACKNOWLEDGMENT

This chapter was adapted from Ultrasonic Inspection in *Metals Handbook Desk Edition*, Second Edition, 1998.

REFERENCES

1. Ultrasonic Inspection, *Metals Handbook Desk Edition*, 2nd ed., ASM International, 1998, p 1282–1290

SELECTED REFERENCES

- L. Cartz, *Nondestructive Testing*, ASM International, 1995
- K.-J. Langenberg, R. Marklein, and K. Mayer, *Ultrasonic Nondestructive Testing of Materials: Theoretical Foundations*, CRC Press, 2012
- *Nondestructive Testing and Quality Control*, Vol 17, *ASM Handbook*, ASM International, 1989
- C.H. Shen, *Ultrasonic and Advanced Methods for Nondestructive Testing and Materials Characterization*, World Scientific, 2007

CHAPTER 12

Inspection of Castings

INSPECTION PROCEDURES FOR CASTINGS are established at the foundry to ensure conformance with customer drawings and documents, which are frequently based on various government, technical society, or commercial specifications. For a foundry to ensure casting quality, inspection procedures must be efficiently directed toward the prevention of imperfections, the detection of unsatisfactory trends, and the conservation of material, all of which ultimately lead to reduction in costs. Inspectors should be able to assess on sight the probable strong and weak points of a casting and know where weaknesses and faults would most likely be found.

The inspection of castings normally involves checking for shape and dimensions, coupled with aided and unaided visual inspection for external discontinuities and surface quality. Chemical analyses and tests for mechanical properties are supplemented by various forms of nondestructive inspection, including leak testing and proof loading, all of which are used to evaluate the soundness of the casting. These inspections add to the cost of the product; therefore, the initial consideration must be to determine the amount of inspection needed to maintain adequate control over quality. In some cases, this may require full inspection of each individual casting, but in other cases sampling procedures may be sufficient.

Inspection Categories

Methods for Determining Surface Quality. Cracks and other imperfections at the surface of a casting can be detected by a number of inspection techniques, including visual inspection, chemical etching, liquid penetrant inspection, eddy current inspection, and magnetic particle inspection, which can also reveal discontinuities situated immediately below the surface. All these inspection methods require clean and relatively smooth surfaces for effective results.

Methods for Detecting Internal Discontinuities. The principal nondestructive methods used for detecting internal discontinuities in castings are radiography, ultrasonic inspection, and eddy current inspection. Of these methods, radiography is the most highly developed technique for detailed inspection; it can provide a pictorial representation of the form and extent of many types of internal discontinuities. Ultrasonic inspection, which is less universally applicable, can give qualitative indications of many discontinuities. It is especially useful in the inspection of castings of fairly simple design, for which the signal pattern can be most reliably interpreted. Ultrasonic inspection can also be used to determine the shape of graphite particles in cast iron.

Eddy current and other closely related electromagnetic methods are used to sort castings for variations in composition, surface hardness, and structure. Infrared thermography (thermal inspection) has also occasionally been proposed as a method for detecting subsurface defects. However, its successful uses have generally been restricted to the detection of larger defects because of the relatively slow rates at which heat can be put into a component and because of the relatively low sensitivity of infrared detectors. Increased use of thermal inspection may occur with the introduction of pulsed video thermography, in which a very short burst of intense heat is directed at the component. The presence of near-surface defects influences the rate at which heat is dissipated from the surface, and temperature variations are detected with a high resolution infrared camera recorded onto videotape and presented as an image on a TV monitor. The method was developed for the detection of small defects in composites and in aerospace turbine engine blades, but some initial results obtained with cast iron test plates have proved promising.

Methods for Dimensional Inspection. A number of techniques are used to determine the dimensional accuracy of castings. These include manual checks with micrometers, manual and automatic gages, coordinate measuring machines, and three-dimensional automatic inspection stations (machine vision systems).

Casting Defects

Poor casting design can interfere with the ability of the foundry to use the best techniques to produce reliable castings. The designer also specifies the quality requirements that ensure that the cast component will perform as desired. Over specification causes needless expense and can be avoided by understanding the effect of discontinuities on casting performance, and the effect of casting design on the tendency for discontinuities to form during the casting process. Important types of casting discontinuities include porosity, inclusions, oxide films, second phases, hot tears, metal penetration, and surface defects. A number of typical casting defects are shown in Fig. 1.

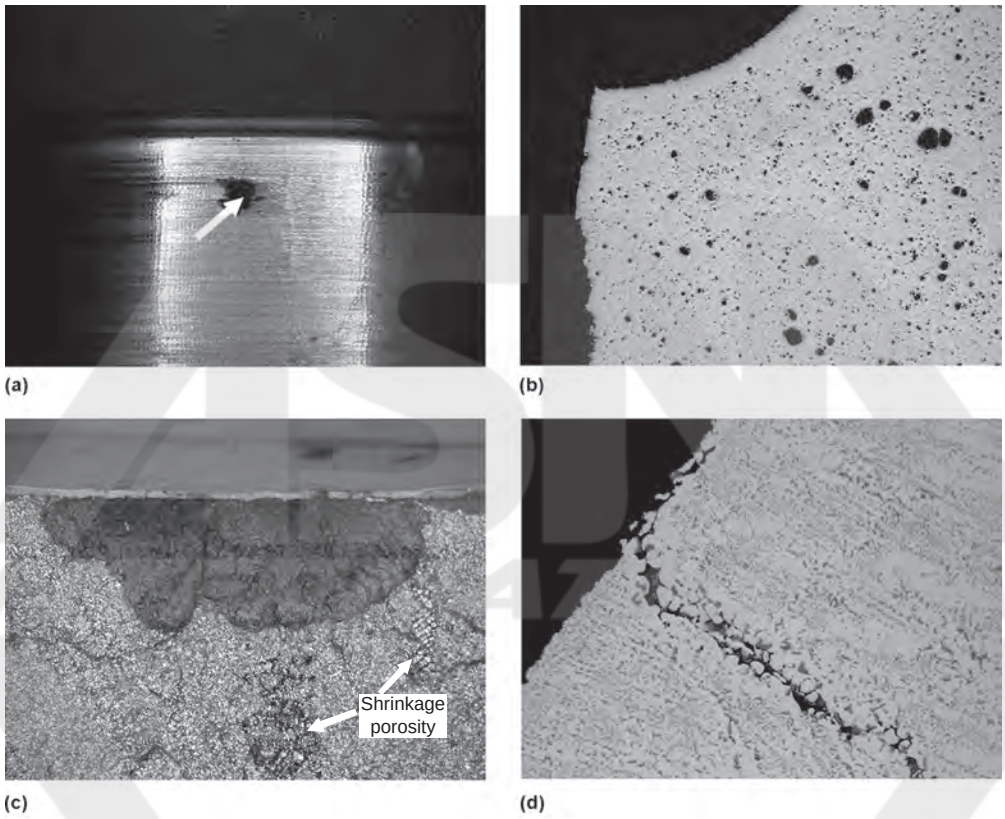


Fig. 1 Typical casting defects. (a) Inclusion (arrow) on machined surface of a casting. (b) Typical micrograph of gas porosity. Original magnification: 100×. (c) Micrograph of low alloy steel shrinkage crack. Original magnification: 7.5×. (d) Optical micrograph of a hot tear in a casting. Original magnification: 200×. Source: Ref 1

Porosity is a common defect in castings and takes many forms. An example of gas porosity is shown in Fig. 1(b). Pores may be connected to the surface, where they can be detected by dye penetrant techniques, or they may be wholly internal, where they require radiographic techniques to detect. Macroporosity refers to pores that are large enough to see with the unaided eye on radiographic inspection, while microporosity refers to pores that are not visible without magnification. Both macroporosity and microporosity are caused by the combined action of metal shrinkage and gas evolution during solidification. It has been shown that nucleation of pores is difficult in the absence of some sort of substrate, such as a nonmetallic inclusion, a grain refiner, or a second phase particle. This is why numerous investigations have shown that clean castings, those castings that are free from inclusions, have fewer pores than castings that contain inclusions.

When the shrinkage and the gas combine to form macroporosity, properties are deleteriously affected. Static properties are reduced at least by

the portion of the cross-sectional area that is taken up with the pores; because there is no metal in the pores, there is no metal to support the load there, and the section acts as though its area was reduced. Because the pores may also cause a stress concentration in the remaining material, static properties may be reduced by more than the percentage of cross-sectional area that is caused by the macroporosity. Dynamic properties are also affected by porosity. A study of aluminum alloys showed that fatigue properties in some were reduced 11% when specimens having x-ray quality equivalent to ASTM E155 level 4 were tested, and that they were reduced 17% when specimens having quality of ASTM E155 level 8 were tested.

Static properties are mostly unaffected by microporosity. Microporosity is found between dendrites, and like macroporosity, is caused by the inability of feed metal to reach the interdendritic areas of the casting where shrinkage is occurring and where gas is being evolved. However, because this type of porosity occurs late in solidification, particularly in long range freezing (mushy freezing) alloys, it is particularly difficult to eliminate. The most effective method is to increase the thermal gradient, often accomplished by increasing the solidification rate, which decreases the length of the mushy zone. However, this technique may be limited by alloy and mold thermal properties, and by casting geometry, that is, the design of the casting.

As long as the micropores are less than 0.2 mm (0.008 in.) in length, there is no effect on dynamic properties; fatigue properties of castings with pores that size or smaller are in the same range as those of castings where no micropores were found. The shape of the micropore is as important as its size, with elongated pores having a greater effect than round pores. Microporosity can be healed by hot isostatic pressing (HIP). In one study comparing HIP and non-HIP samples, no difference was found in fatigue lives of HIP and non-HIP samples. However, the HIP samples showed a lower crack growth rate than non-HIP samples. In another study, HIP improved fatigue crack growth resistance only close to threshold levels. The design of the casting directly affects its tendency to solidify in a progressive manner, thereby affecting both the quality and the price of the cast component. Porosity and casting costs are minimized in casting designs that emphasize progressive solidification toward a gate or riser, tapered walls, and the avoidance of hot spots.

Inclusions are nonmetallic particles that are found in the casting. They may form during solidification as some elements (notably manganese and sulfur in steel) precipitate from solution in the liquid. More frequently, they are formed before solidification begins. The former are sometimes called indigenous inclusions, and the latter are called exogenous inclusions. Inclusions are ceramic phases that have little ductility. A crack may form in the inclusion and propagate from the inclusion into the metal, or a crack may form at the interface between the metal and the inclusion. In

addition, because inclusions and the metal have different coefficients of thermal expansion, thermally induced stresses may appear in the metal surrounding the inclusions during solidification. As a result, the inclusions act as a stress concentration and reduce dynamic properties. As in the case of microporosity, the size of the inclusion and its location determines its effect. Small inclusions that are located well within the center of the cross section of the casting have little effect, whereas larger inclusions and those located near the surface of the casting may be particularly detrimental to properties. Inclusions may also be a problem when machining surfaces (Fig. 1a), causing excessive tool wear and tool breakage.

Exogenous inclusions are mostly oxides or mixtures of oxides and are primarily slag or dross particles, which are the oxides that result when the metal reacts with oxygen in the air during melting. These are removed from the melt before pouring by filtration. Most inclusions found in steel castings arise from the oxidation of metal during the pouring operation. This is known as reoxidation, and takes place when the turbulent flow of the metal in the gating system causes the metal to break up into small droplets, which then react with the oxygen in the air in the gating system or casting cavity to form oxides. Metal casters use computer analysis of gating systems to indicate when reoxidation can be expected in a gating system and to eliminate it. However, casting designs that require molten metal to "jet" through a section of the casting to fill other sections will recreate the inclusions and should be avoided.

Oxide films are similar to inclusions and have been found to reduce casting properties. These form on the surface of the molten metal as it fills the mold. If the surface film is trapped within the casting instead of being carried into a riser, it forms a linear discontinuity and an obvious site for crack initiation. It has been shown that elimination of oxide films, in addition to substantially improving static properties, results in up to a five-fold improvement of fatigue life in axial tension-tension tests.

Oxide films are of particular concern in nonferrous castings, although they also must be controlled in steel and stainless steel castings. If the film folds over on itself as a result of turbulent flow or *waterfalling* when molten metal falls to a lower level in the casting during mold filling, the effects are particularly damaging. Casting design influences how the metal fills the mold, and features of the design that require the metal to fall from one level to another while the mold is filling should be avoided so that waterfalls are eliminated. Oxide films are avoided by filling the casting from the bottom, in a controlled manner, by pumping the metal into the mold using pneumatic or electromagnetic pumps.

Second phases, which form during solidification, can also nucleate cracks if they have the proper size and morphology. An example is aluminum-silicon alloys, where the silicon eutectic is present as large platelets, which can nucleate cracks, and along which cracks propagate. The size of these platelets may be significantly reduced by modifying the alloy with

additions of sodium or strontium. However, such additions increase the size of micropores, and for this reason, many foundrymen rely on accelerated solidification of the casting to refine the silicon. As the solidification rates increase, the structure is refined in thin sections. Heavy sections are to be avoided if a fine structure is desired.

Hot tears form when casting sections are constrained by the mold from shrinking as they cool near the end of solidification. These discontinuities are fairly large and are most often weld repaired. If not repaired, their effect is not readily predictable. While generally they are detrimental to casting properties, under some circumstances they do not affect them. Hot tears (Fig. 1d) are caused by a combination of factors, including alloy type, metal cleanliness, and mold and core hardness. However, poor casting design is the primary cause. Castings should be designed so that solidifying sections are not subjected to tensile forces caused by shrinkage during solidification, as the solidifying alloy has little strength before it solidifies.

Metal Penetration. Molten metal may penetrate the surface of the mold, forming a rough surface or, in extreme cases, actually becoming intimately mixed with the sand in the mold. In iron castings, this is normally the result of the combination of metallostatic head (the pressure exerted on the molten iron at the bottom of the mold by the weight of the metal on top of it) and the surface tension relationship between the liquid iron and molding materials. In cast iron, it is also frequently the result of the expansion of graphite at the end of solidification, forcing liquid metal into the mold if the casting is not properly designed with a tapered wall to promote directional solidification and avoid hot spots.

Surface Defects. Surface finish is also an important property. Surface discontinuities affect fatigue life, and obviously smoother surfaces are superior to rough surfaces. Designers should be certain that fatigue data used in design calculations has been taken from as-cast surfaces rather than machined surfaces, as most surfaces on castings where stress concentrations might be expected are not machined. Surface finish in castings is controlled by the application of coatings to the mold as well as proper selection of mold materials. Metal mold casting processes generally produce better surface finishes than sand casting processes.

Design and Service Considerations. The existence of casting discontinuities does not, in and of itself, indicate that casting performance in service will be affected. Equally important are the size, location, and distribution of these discontinuities. Those discontinuities that are small and located near the center of the casting have little effect, while those located at or near the surface of the casting are usually damaging. Clustered discontinuities and those that occur in a regular array have a greater effect on properties than those that are isolated and randomly distributed. In specifying acceptable levels of discontinuities, such as microporosity and inclusion sizes and distribution, the designer should determine the critical

flaw size that will deleteriously affect performance in service. This permits the foundry to design a casting practice that will eliminate such discontinuities at minimum cost.

Common Inspection Procedures

The inspection of castings is most often limited to visual and dimensional inspections, weight testing, and hardness testing. However, for castings that are to be used in critical applications, such as automotive or aerospace components, additional methods of nondestructive inspection are used to determine and to control casting quality.

Visual inspection of each casting ensures that none of its features has been omitted or malformed by molding errors, short running, or mistakes in cleaning. Most surface defects and roughness can be observed at this stage.

Initial sample castings from new pattern equipment should be carefully inspected for obvious defects. Liquid penetrant inspection can be used to detect surface defects. Such casting imperfections as shrinks, cracks, blows, or dross usually indicate the need for adjustment in the gating or foundry techniques. If the casting appears to be satisfactory based on visual inspection, internal quality can be checked by radiographic and ultrasonic inspection.

The first visual inspection operation on the production casting is usually performed immediately after shakeout or knockout of the casting, ensuring that major visible imperfections are detected as quickly as possible. This information, promptly relayed to the foundry, permits early corrective action to be taken with a minimum of scrap loss. The size and complexity of some sand castings requires that the gates and risers be removed to permit proper inspection of the casting.

Many castings that contain numerous internal cores or have close dimensional tolerances require a rapid, but fairly accurate check of critical wall dimensions. In some cases, an indicating type caliper gage is suitable for this work, and special types are available for casting shapes that do not lend themselves to the standard types. Ultrasonic inspection is also used to determine wall thickness in such components as cored turbine blades made by investment casting. New developments in visual inspection procedures for examining component appearance are mainly based on vision systems that use electronic cameras coupled to computer-assisted image processing systems.

With the development of high sensitivity cameras having exposure times of 1/1000 s, components can be inspected on moving belts. Flexibility for examining three-dimensional components can be achieved with an array of cameras multiplexed to a common image processor or with a computer controlled camera scanning system. Such systems have been successfully applied to the inspection of printed circuit boards in the elec-

tronics industry and subassemblies in automobile manufacture. These tests usually operate on a go/no-go basis; either the assembly is complete with connections correctly made or it is not correct. This is a far easier task than evaluating casting quality.

Studies that have been carried out to assess the possibility of extending such methods to iron castings have not given encouraging results. Contrast between defective and nondefective areas is low, illumination is critical, and consistent standards of inspection are difficult to maintain because of differences in reflectivity of the casting surfaces depending on whether or not they have been recently shot blasted. Even the simple task of identifying castings to determine their type is best carried out by examining their back lighted silhouette, and this provides no advantage in examining their quality.

Dimensional Inspection. Consistency of dimensions is an inescapable requirement of premium quality castings supplied as near net shape components on which subsequent high speed machining operations are to be carried out. Customers will not accept increased machining costs due to inconsistencies in dimensions nor will they tolerate damage to flexible machining systems or transfer times resulting from poor control and inspection in foundries.

Variations in dimensions represent one of the most common complaints with regard to the machinability of iron castings. Prevention is within the control of foundries. Differences in pattern size when using multipattern plates can be virtually eliminated by the use of computer aided design and manufacturing methods and computer numerical control machines in patternmaking. Better process control and pouring methods can eliminate variations in dimensions due to changes in metal composition or feeding methods. Variations in mold rigidity, caused by inadequate compaction with green sand, or the use of cold sand or insufficient curing times with cold setting systems, which cause casting dimensions to fall outside the preset tolerance limits, can be greatly reduced by good molding and coremaking practices.

Because the dimensions and weight of iron castings are directly related to their soundness and are dependent on mold rigidity, the measurement of size or weight provides a simple test for checking casting integrity and for monitoring the consistency of the moldmaking process.

Casting dimensions are usually checked with dial gages, vernier calipers, micrometers, or vertical height gages, which may be hand held or incorporated into acceptance fixtures. Wall thickness measurements can be made with small handheld ultrasonic thickness gages. Under ideal conditions, the accuracy of these instruments is claimed to be ±0.01 mm (±0.004 in.), but this is rarely achieved in practice because the surfaces are not parallel and are not machined. Instruments are available that display variations in thickness from some preset standard and provide a digital readout and a permanent record of results for statistical analysis.

The use of measuring systems employing capacitance, electrical contact, or linear displacement transducers are capable of high accuracy and the output can be linked directly to microcomputers for data recording and statistical analysis to meet the requirements of statistical process control.

Laser methods of measurement using beam displacement or time-lapse techniques are available for use in machine shops where accurate measurement is required for control of automatic machining processes. At present, they are generally not well suited for measuring castings, because of their high cost and because it is difficult to make precise measurements on components having a complex shape with curved or as-cast surfaces. As these laser methods become more widely used in other industries, lower cost systems will become available.

Weight Testing. Many intricately cored castings are extremely difficult to measure accurately, particularly the internal sections. It is important to ensure that these sections are correct in thickness for three main reasons:

- There should be no additional weight that would make the finished product heavier than permissible
- Sections must not be thinner than designed so as not to decrease the strength of the casting
- If hollow cavities have been reduced in area by increasing the metal thickness of the sections, any flow of liquid or gases is reduced

A ready means of testing for these discrepancies is by accurately weighing each casting or by measuring the displacement caused by immersing the casting in a liquid filled measuring vessel. In certain cases in which extreme accuracy is demanded, a tolerance of only ±1% of a given weight may be allowed.

Hardness testing is often used to verify the effectiveness of the heat treatment applied to actual castings. Its general correlation with the tensile strength of many ferrous alloys allows a rough prediction of tensile strength to be made.

The Brinell hardness test is most frequently used for casting alloys. Combination of a large diameter ball (5 or 10 mm or 0.2 or 0.4 in.) and heavy load (500 to 3000 kgf or 1100 to 6600 lb) is preferred for the most effective representation because a deep impression minimizes the influence of the immediate surface layer and of the relatively coarse microstructure. The Brinell hardness test is unsuitable for use at high hardness levels (above 600 HB), because distortion of the ball indenter can affect the shape of the indentation.

Either the Rockwell or the Vickers (136° diamond pyramid) hardness test is used for alloys of extreme hardness or for high quality and precision castings in which the large Brinell indentation cannot be tolerated. Because of the very small indentations produced in Rockwell and Vickers tests, which use loads of 150 kg (330 lb) or less, results must be based on

the average of a number of determinations. Portable hardness testers or ultrasonic microhardness testers can be used on large castings that cannot be placed on the platform of a bench type machine.

The hardness of ferrous castings can be determined from the sonic velocity of the metal if all other test conditions remain constant. This has been demonstrated on chilled rolls in determining the average hardness of the core.

Computer-Aided Dimensional Inspection

The use of computer equipment in foundry inspection operations is finding more acceptance as the power and utility of available hardware and software increase. The computerization of operations can reduce the manhours required for inspection tasks, can increase accuracy, and can allow the analysis of data in ways that are not possible or practical with manual operations. Perhaps the best example of this, given the currently available equipment, is the application of computer technology to the dimensional inspection of castings.

Importance of Dimensional Inspection. One of the most critical determinants of casting quality in the eyes of the casting buyer is dimensional accuracy. Parts that are within dimensional tolerances, given the absence of other casting defects, can be machined, assembled, and used for their intended functions with minimal testing and inspection costs. Major casting buyers are therefore demanding statistical evidence that dimensional tolerances are being maintained. In addition, the statistical analysis of in-house processes has been demonstrated to be effective in keeping those processes under control, thus reducing scrap and rework costs.

The application of computer equipment to the collection and analysis of dimensional inspection data can increase the amount of inspection that can be performed and decrease the time required to record and analyze the results. This furnishes control information for making adjustments to tooling on the foundry floor and statistical information for reporting to customers on the dimensional accuracy of parts.

Typical Equipment. A typical installation for the dimensional inspection of castings consists of an electronic coordinate measuring machine, a microcomputer interfaced to the coordinate measuring machine controller with a data transfer cable, and a software system for the microcomputer (Fig. 2). The software system should be capable of controlling the functions and memory storage of the coordinate measuring machine, as well as recalling and analyzing the data it collects. The software serves as the main control element for the dimensional inspection and statistical reporting of results. Such software can be purchased or, if the expertise is available, developed in-house for highly specialized requirements.

Coordinate measuring machines typically record dimensions along three-axes from data points specified by the user. Depending on the so-

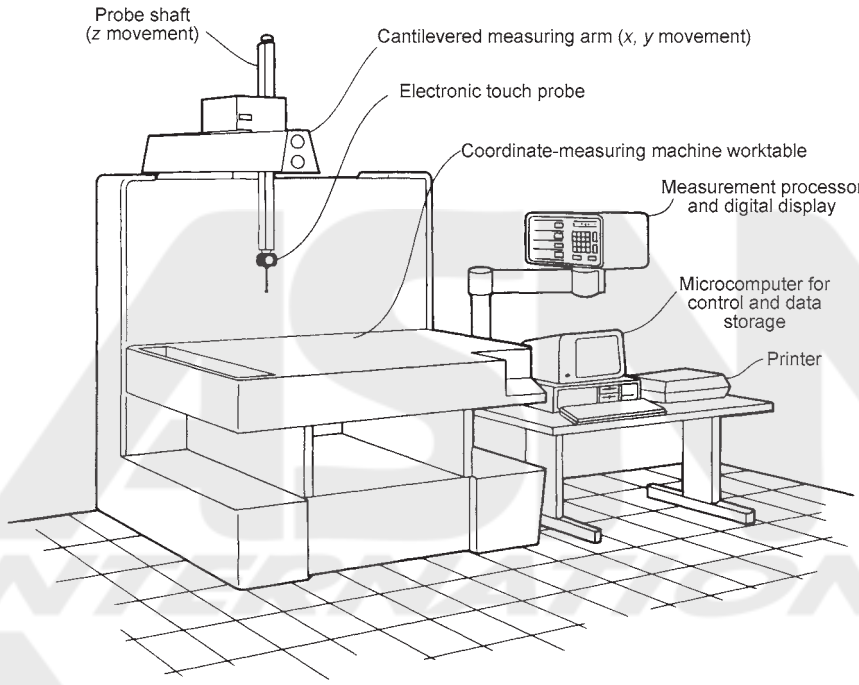


Fig. 2 Equipment used in a typical installation for the computer-aided dimensional inspection of castings showing a coordinate measuring machine and microcomputer. Source: Ref 2

phistication of the controller, such functions as center and diameter finds for circular features and the electronic rotation of measurement planes can be performed. Complex geometric constructions, such as the intersection points of lines and planes and out-of-roundness measurements, are typically off-loaded for calculation into the microcomputer. The contact probe of the coordinate measuring machine can be manipulated manually, or in the case of direct computer controlled machines, the probe can be driven by servomotors to perform the part measurement with little operator intervention.

The Measurement Process. Figure 3 illustrates the general procedure in applying semiautomatic dimensional inspection to a given part. The first step is to identify the critical part dimensions that are to be measured and tracked. Nominal dimensions and tolerances are usually taken from the customer's specifications and blueprints. Dimensions that are useful in controlling the foundry process can also be selected. A data-base file, including a description and tolerance limits for each dimension to be checked, is then created using the microcomputer software system.

The next step in the setup process is to develop a set of instructions for measuring the part with the coordinate measuring machine. The instructions consist of commands that the coordinate measuring machine uses to

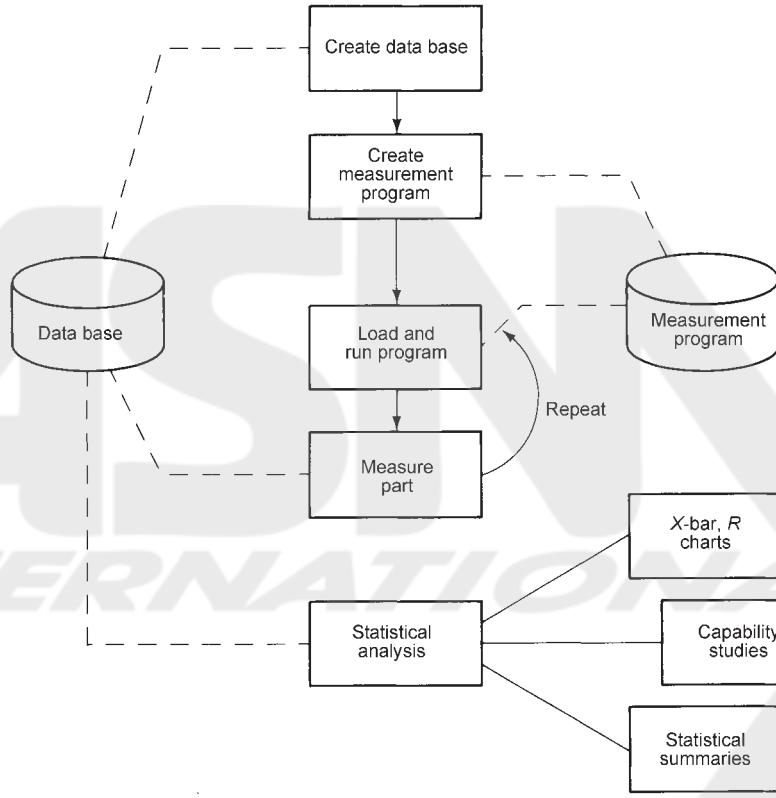


Fig. 3 Flowchart showing typical sequence of operations for computer-aided dimensional inspection. Source: Ref 2

establish reference planes and to measure such features as the center points of circular holes.

This measurement program can be entered in either of two ways. In the first method, the operator simply types in a list of commands that he wants the coordinate measuring machine to execute and that give the required dimensions as defined in the part data base. The second method uses a teach mode. The operator actually places a part on the worktable of the coordinate measuring machine and checks it in the proper sequence, while the computer monitors the process and stores the sequence of commands used. In either case, the result is a measurement program stored on the microcomputer that defines in precise detail how the part is to be measured. Special commands can also be included in the measurement program to display operator instructions on the computer screen while the part is being measured.

In developing the measurement program, consideration must be given to the particular requirements of the part being measured. Customer prints will normally show datum planes from which measurements are to be made. When using a cast surface to establish a datum plane, it is good

practice to probe a number of points on the surface and to allow the computer to establish a best fit plane through the points. Similarly, the center points of cast holes can best be found by probing multiple points around the circumference of the hole. Machined features can generally be measured with fewer probe contacts. When measuring complex castings, maximum use should be made of the ability of the coordinate measuring machine to electronically rotate measurement planes without physically moving the part; unclamping and turning a part will lower the accuracy of the overall layout.

Once the setup process is complete, the dimensional inspection of parts from the foundry begins. Based on statistical considerations, a sampling procedure and frequency must be developed. Parts are then selected at random from the process according to the agreed upon frequency. The parts are brought to the coordinate measuring machine, and the operator calls up the measurement program for that part and executes it. As the part is measured, the dimensions are sent from the coordinate measuring machine to the data base on the microcomputer. Once the measurement process is completed, information such as mold number, shift, date, and serial number should be entered by the operator so that this particular set of dimensions can be identified later. A layout report can then be generated to show how well the measured part checked out relative to the specified dimensions and tolerances. Figure 4 shows a sample report in the form of a

LAYOUT REPORT

FILE: FP5.DAT DESCRIPTION: FUEL PUMP MOLD 5 REC. NO: 40
MONTH/DAY: 02/20 YEAR: 87 SHIFT : 1 MOLD : 5

	DESCRIPTION	ACT	MEAN	- DEV	0	+ DEV
1	DIM A	2.813	2.790		|----------	
2	DIM B	2.501	2.500			
3	DIM C	0.547	0.560	------		
4	DIM D	4.453	4.440		|------	
5	DIM E	1.622	1.620		|-	
6	DIM F	1.873	1.880	---		
7	DIM G	1.658	1.670	----		
8	DIM H	1.576	1.560		|-------	
9	DIM I	1.358	1.350		|---	
10	DIM J	1.657	1.670	-----		
11	DIM K	3.456	3.435		|--------	
12	DIM L	0.750	0.750				
13	DIM MX	2.791	2.790				
14	DIM MY	0.325	0.370	--------|		
15	SNP RG A	2.156	2.190	--------|		
16	SNP RG B	3.237	3.235		|-	
17	MLD HF R	0.658	0.655		|-	
18	MLD HF L	3.986	3.970		|-----	
19	WALL A	0.206	0.214	------|		
20	WALL B	0.159	0.153		|----	
21	WALL C	0.237	0.228		|-----	
22	OAH MAX	3.838	3.770		|----------*******	
23	OAH MIN	3.802	3.770		|--------	

Fig. 4 Example layout report showing all dimensions measured on a single casting, with visual indication of deviations from print mean. Note out-of-tolerance condition indicated by asterisks. Source: Ref 2

bar graph, in which any deviation from print tolerance appears as a line of dashes to the left or right of center. A deviation outside of tolerance limits displays asterisks to flag its condition. Such a report is useful in that it gives a quick visual indication of the measurement of the casting.

Statistical analysis permits the mathematical prediction of the characteristics of all the parts produced by measuring only a sample of those parts. All processes are subject to some amount of natural variation; in most processes, this variation follows a normal distribution (the familiar bell shaped curve) when the probability of occurrence is plotted against the range of possible values. The standard deviation, a measure of the distance from center on the probability curve, is the principal means of expressing the range of measured values. For example, a spread of six standard deviations (plus or minus three standard deviations on either side of the measured mean) represents the range within which one would expect to find 99.73% of observed measurements for a normal process. This allows the natural variation inherent in the process to be quantified.

Control Charts. With statistical software incorporated into the microcomputer system, the results of numerous measurements of the same part can be analyzed to determine, first, how well the process is staying in control, that is, whether the natural variations occurring in a given measurement are within control limits and whether any identifiable trends are occurring. This is done by using a control chart (Fig. 5), which displays the average values and ranges of groups of measurements plotted against time. Single value charts with a moving range can also be helpful. The control limits can also be calculated and displayed. With the computer, this type of graph can be generated within seconds. Analysis of the graph may show a developing trend that can be corrected by adjusting the tooling before out-of-tolerance parts are made.

Statistical Summary Report. The second type of analysis shows the capability of the process, which is how the range of natural variation (as measured by a specified multiple of the standard deviation) compares with the tolerance range specified for a given dimension. An example of a useful report of this type is shown in Fig. 6. This information is of great interest both to the customer and the process engineer, because it indicates whether or not the process being used to produce the part can hold the dimensions within the required tolerance limits. The user must be aware that different methods of capability analysis are used by different casting buyers, so the software should be flexible enough to accommodate the various methods of calculation that might be required.

Histograms. An alternative method of assessing capability involves the use of a histogram, or frequency plot. This is a graph that plots the number of occurrences within successive, equally spaced ranges of a given measured dimension. Figure 7 shows an example output report of this type. A graph such as this, which has superimposed upon it the tolerance limits for the dimension being analyzed, allows a quick, qualitative evaluation of

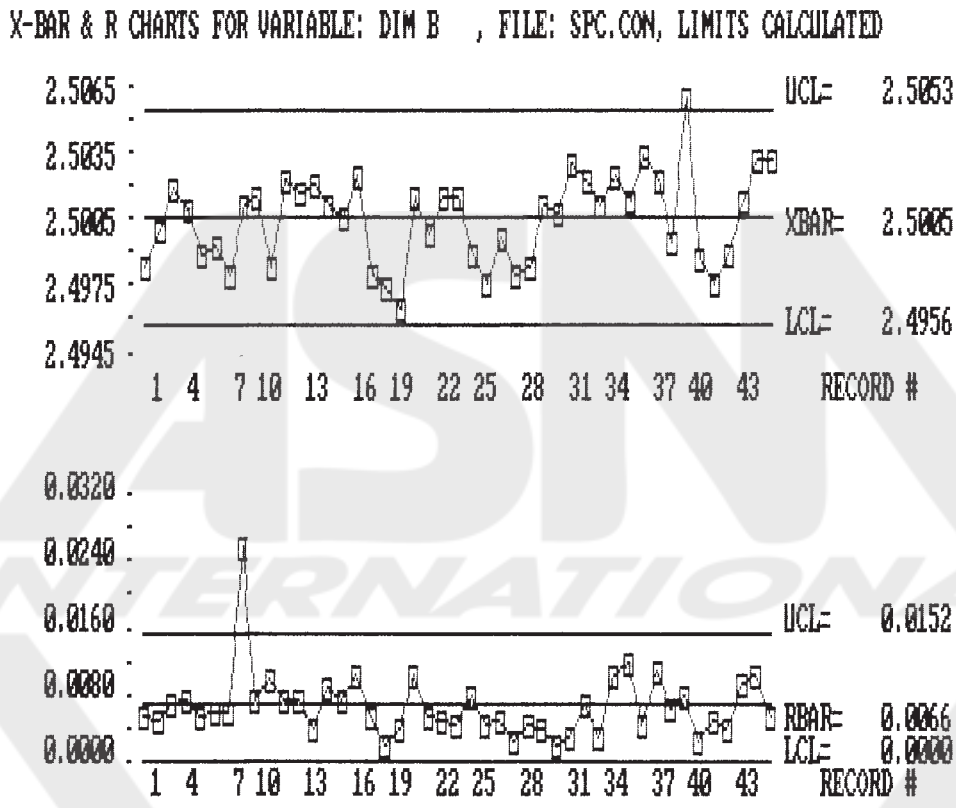


Fig. 5 Control chart with average of groups of measurements (*X* values) plotted above and ranges within the groups (*R*) plotted below. Control limits have been calculated and placed on the chart by the computer. Source: Ref 2

the variation and capability of the process. It also allows the normality of the process to be judged through comparison with the expected bell-shaped curve of a normal process.

Other Applications for Computer-Aided Inspection. The general sequence described for semiautomatic dimensional inspection can be applied to a number of other inspection criteria. Examples would include pressure testing or defect detection by electronic vision systems. The statistical analysis of scrap by defect types is also very helpful in identifying problem areas. In some cases, direct data input to a computer may not be feasible, but the benefits of entering data manually into a statistical analysis program should not be overlooked. The computer allows rapid analysis of large amounts of data so that statistically significant trends can be detected and proper attention paid to appropriate areas for improvement. The benefits and costs of each anticipated application of automation to a particular situation, as well as the feasibility of applying state-of-the-art equipment need to be studied as thoroughly as possible prior to implementation.

Statistical Summary Report Page 1
FILE: FP5.DAT DESCRIPTION: FUEL PUMP MOLD 5
For SHIFT : ALL MOLD : ALL Start Rec: 1 End Rec: 68

		Actual Mean	Spec.	Spec.- Mean	Tol.	Sigma	6 Sigma	Process Capab
1	DIM A	2.8025	2.7900	-0.0125	0.0460	0.0070	0.0419	1.0976
2	DIM B	2.5017	2.5000	-0.0017	0.0460	0.0059	0.0352	1.3072
3	DIM C	0.5537	0.5600	0.0063	0.0460	0.0051	0.0304	1.5152
4	DIM D	4.4517	4.4400	-0.0117	0.0460	0.0083	0.0496	0.9281
5	DIM E	1.6182	1.6200	0.0018	0.0460	0.0074	0.0447	1.0300
6	DIM F	1.8665	1.8800	0.0135	0.0460	0.0218	0.1305	0.3524
7	DIM G	1.6607	1.6700	0.0093	0.0460	0.0088	0.0526	0.8740
8	DIM H	1.5671	1.5600	-0.0071	0.0460	0.0097	0.0580	0.7931
9	DIM I	1.3621	1.3500	-0.0121	0.0460	0.0044	0.0265	1.7354
10	DIM J	1.6597	1.6700	0.0103	0.0460	0.0106	0.0639	0.7201
11	DIM K	3.4469	3.4350	-0.0119	0.0460	0.0093	0.0560	0.8212
12	DIM L	0.7479	0.7500	0.0021	0.0460	0.0047	0.0282	1.6326
13	DIM MX	2.7821	2.7900	0.0079	0.0900	0.0102	0.0609	1.4771
14	DIM MY	0.3338	0.3700	0.0362	0.1000	0.0108	0.0648	1.5433
15	SNP RG A	2.1586	2.1900	0.0314	0.0900	0.0108	0.0648	1.3899
16	SNP RG B	3.2267	3.2350	0.0083	0.0900	0.0114	0.0683	1.3170
17	MLD HF R	0.6615	0.6550	-0.0065	0.0600	0.0036	0.0216	2.7834
18	MLD HF L	3.9807	3.9700	-0.0107	0.0600	0.0083	0.0498	1.2055
19	WALL A	0.2133	0.2140	0.0007	0.0280	0.0084	0.0502	0.5573
20	WALL B	0.1541	0.1530	-0.0011	0.0300	0.0097	0.0581	0.5161
21	WALL C	0.2286	0.2280	-0.0006	0.0400	0.0073	0.0440	0.9097
22	OAH MAX	3.7892	3.7700	-0.0192	0.0800	0.0352	0.2110	0.3791
23	OAH MIN	3.7589	3.7700	0.0111	0.0800	0.0244	0.1464	0.5465

68 Parts included in analysis
Actual Sigma calculated

$$\text{Process Capability} = \frac{\text{Blueprint Tolerance}}{6 \times \text{Sigma}}$$

Fig. 6 Statistical summary report showing the mean of measured observations; the blueprint specification for the mean; the difference between specified and measured means; the tolerance; the standard deviation of the measured dimensions; and the capability of the process. These calculations are performed for all measured dimensions on the part. Source: Ref 2

Liquid Penetrant Inspection

Liquid penetrants will highlight surface defects so that detection is more certain. Liquid penetrant inspection should not be confined to as-cast surfaces. For example, it is not unusual for castings of various alloys to exhibit cracks (frequently intergranular) on machined surfaces. A pattern of cracks of this type may be the result of intergranular cracking throughout the material because of an error in composition or heat treatment, or the cracks may be on the surface only as a result of machining or grinding. Surface cracking may result from insufficient machining allowance, which does not allow for complete removal of imperfections produced on the as-cast surface, or it may result from faulty machining techniques. If imperfections of this type are detected by visual inspection, liquid penetrant inspection will show the full extent of such imperfections, will give some indication of the depth and size of the defect below the surface by the amount of penetrant absorbed, and will indicate whether

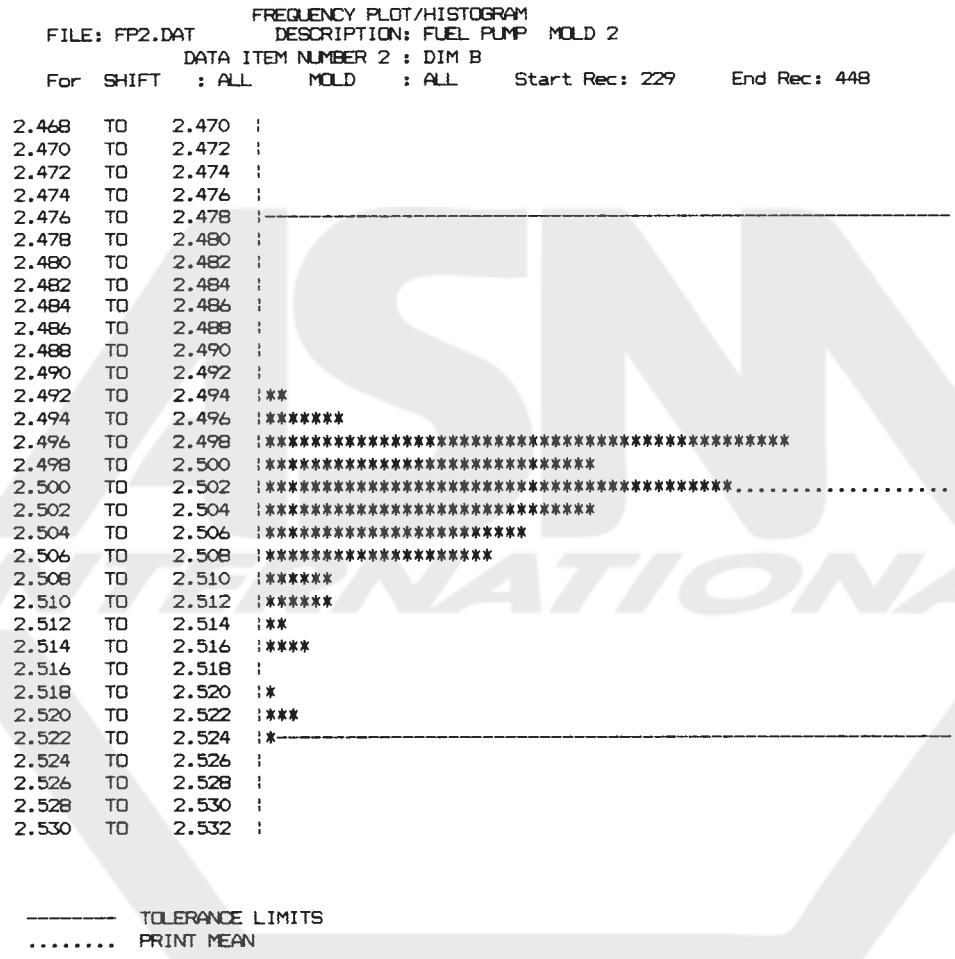


Fig. 7 Frequency plot for one measured dimension showing the distribution of the measurements. Source: Ref 2

cracking is present throughout the section. Liquid penetrant inspection is sometimes used in conjunction with another nondestructive method.

Magnetic Particle Inspection

Magnetic particle inspection is a highly effective and sensitive technique for revealing cracks and similar defects at or just beneath the surface of castings made of ferromagnetic metals. The capability of detecting discontinuities just beneath the surface is important because such cleaning methods as shot or abrasive blasting tend to close a surface break that might go undetected in visual or liquid penetrant inspection.

Equipment for magnetic particle inspection uses direct or alternating current to generate the necessary magnetic fields. The current can be applied in a variety of ways to control the direction and magnitude of the

magnetic field. In one method of magnetization, a heavy current is passed directly through the casting placed between two solid contacts. The induced magnetic field then runs in the transverse or circumferential direction, producing conditions favorable to the detection of longitudinally oriented defects. A coil encircling the casting will induce a magnetic field that runs in the longitudinal direction, producing conditions favorable to the detection of circumferentially (or transversely) oriented defects. Alternatively, a longitudinal magnetic field can be conveniently generated by passing current through a flexible cable conductor, which can be coiled around any metal section. This method is particularly adaptable to castings of irregular shape. Circumferential magnetic fields can be induced in hollow cylindrical castings by using an axially disposed central conductor threaded through the casting.

Small castings can be magnetic particle inspected directly on bench type equipment that incorporates both coils and solid contacts. Critical regions of larger castings can be inspected by the use of yokes, coils, or contact probes carried on flexible cables connected to the source of current; this setup enables most regions of castings to be inspected.

Eddy Current Inspection

Eddy current methods of inspection are effective with both ferromagnetic and nonferromagnetic castings. Eddy current methods are not as sensitive to small, open defects as liquid penetrant or magnetic particle methods are. Because of the skin effect, eddy current inspection is generally restricted to depths less than 6 mm (¼ in.). The results of inspecting ferromagnetic materials can be obscured by changes in the magnetic permeability of the workpiece. Changes in temperature must be avoided to prevent erroneous results if electrical conductivity or other properties, including metallurgical properties, are being determined.

Applications of eddy current and electromagnetic methods of inspection to castings can be divided into the following three categories:

- Detecting near-surface flaws such as cracks, voids, inclusions, blowholes, and pinholes (eddy current inspection)
- Sorting according to alloy, temper, electrical conductivity, hardness, and other metallurgical factors (primarily electromagnetic inspection)
- Gaging according to size, shape, plating thickness, or insulation thickness (eddy current or electromagnetic inspection)

Radiographic Inspection

Internal flaws, such as gas entrapment or nonmetallic inclusions, have a direct effect on radiation attenuation. These flaws create variations in material thickness, resulting in localized dark or light spots on the image.

Sensitivity or the ability to detect flaws, of radiographic inspection, depends on close control of the inspection technique, including the geometric relationships among the point of x-ray emission, the casting, and the x-ray imaging plane. The smallest detectable variation in metal thickness lies between 0.5 and 2.0% of the total section thickness. Narrow flaws, such as cracks, must lie in a plane approximately parallel to the emergent x-ray beam to be imaged; this requires multiple exposures for x-ray film techniques and a remote control parts manipulator for a real-time system.

Real-time systems have eliminated the need for multiple exposures of the same casting by dynamically inspecting parts on a manipulator, with the capability of changing the x-ray energy for changes in total material thickness. These capabilities have significantly improved productivity and have reduced costs, thus enabling higher percentages of castings to be inspected and providing instant feedback after repair procedures. Real-time digital radiography images of automotive components are shown in Fig. 8 and 9.

Several advances have been made to assist the industrial radiographer. These include the computerization of the radiographic standard shooting sketch, which graphically shows areas to be x-rayed and the viewing direction or angle at which the shot is to be taken, and the development of microprocessor controlled x-ray systems capable of storing different x-ray

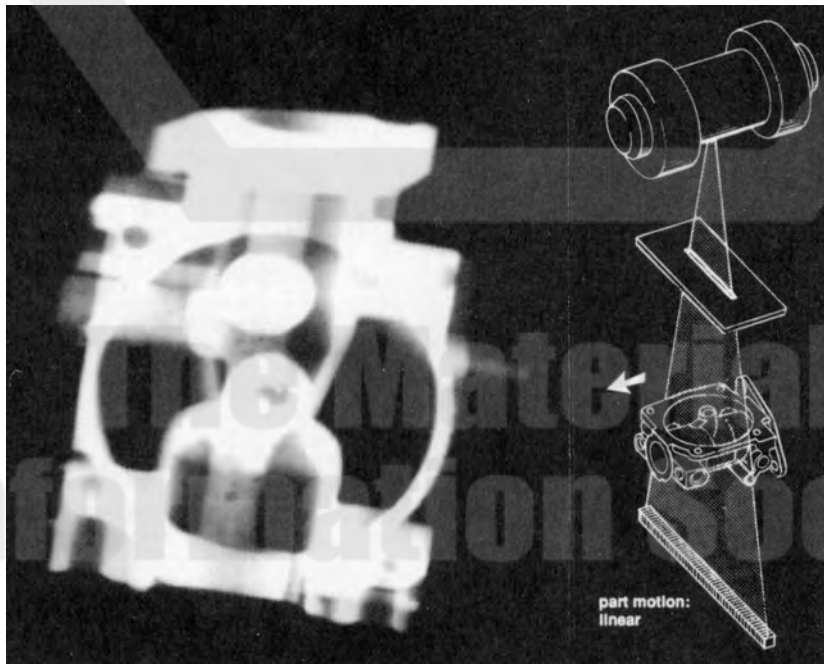


Fig. 8 Digital radiography image of a die cast aluminum carburetor. Porosity appears as dark spots in the area of the center bore, through the vertical center of the image. Courtesy of B.G. Isaacson, Bio-Imaging Research, Inc.

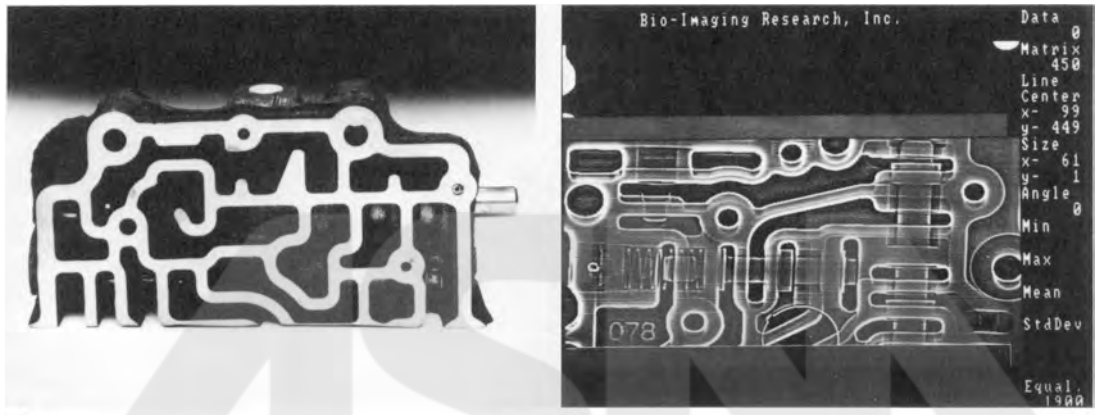


Fig. 9 Evaluation of cast transmission housing assembly. (a) Photograph of cast part. (b) Digital radiography image used to verify the steel spring pin and shuttle valve assembly through material thicknesses ranging from 3 mm (1/8 in.) in the channels to 25 mm (1 in.) in the rib sections of the casting. Courtesy of B.G. Isaacson, Bio-Imaging Research, Inc.

exposure parameters for rapid retrieval and automatic warm-up of the system prior to use. The advent of digital image processing systems and microfocus x-ray sources (near point source) producing energies capable of penetrating thick material sections have made real-time inspection capable of producing images equal to, and in some cases superior to, x-ray film images by employing geometric relations previously unattainable with macrofocus x-ray systems. The near point source of the microfocus x-ray system virtually eliminates the edge unsharpness associated with larger focus devices.

Digital image processing can be used to enhance imagery by multiple video frame integration and averaging techniques that improve the signal-to-noise ratio of the image. This enables the radiographer to digitally adjust the contrast of the image and to perform various edge enhancements to increase the clarity of many linear indications (Fig. 10).

Interpretation of the radiographic image requires a skilled specialist who can establish the correct method of exposing the castings with regard to x-ray energies, geometric relationships, casting orientation, and can take all of these factors into account to achieve an acceptable, interpretable image. Interpretation of the image must be performed to establish standards in the form of written or photographic instructions. The inspector must also be capable of determining if the localized indication is a spurious indication, a film artifact, a video aberration, or a surface irregularity.

Computed tomography, also known as computerized axial tomography (or CAT scanning), is a more sophisticated x-ray imaging technique originally developed for medical diagnostic use. It is the complete reconstruction by computer of a tomographic plane, or slice of an object. A col-

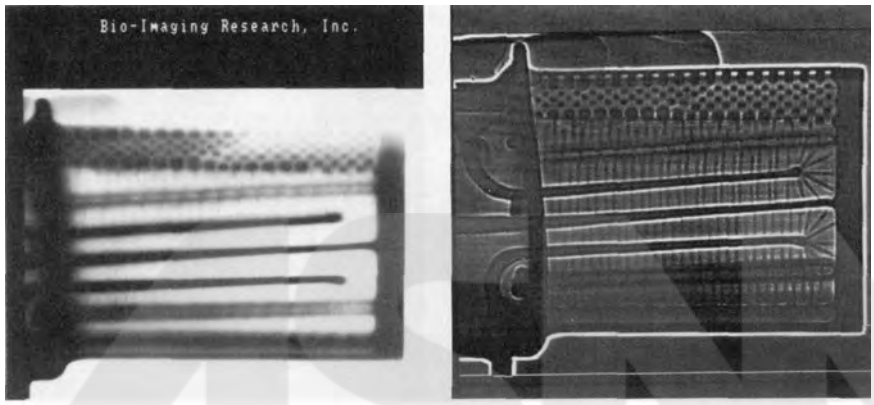


Fig. 10 Digital radiography images of an investment cast jet engine turbine blade showing detail through a wide range in material thickness. The trailing edge of the blade (along the top of the image) is 2 mm (0.080 in.) thick, the root section of the blade (to the far left in the image) is 19 mm (0.75 in.) thick, and the shelf area (to the right of the root section) is 25 mm (1 in.) thick. The image shown in (a) is unprocessed; the image in (b) is processed to subdue the background and to enhance edges and internal features. Courtesy of B.G. Isaacson, Bio-Imaging Research, Inc.

limated (fan shaped) x-ray beam is passed through a section of the part and is intercepted by a detector on the other side. The part is rotated slightly, and a new set of measurements is made; this process is repeated until the part has been rotated 180°.

The resulting image of the slice (tomogram) is formed by computer calculations based on electronic measurement (digital sampling) of the radiation transmitted through the object along different paths during the rotating scan. The data thus accumulated are used to compute the densities of each point in the cross-section, enabling the computer to reconstruct a two-dimensional visual image of the slice. The shapes of internal features are determined by their computed densities. After one slice is produced through a complete rotation either the part or the radiation source and detector can be moved, and a three-dimensional image built up through the scanning of successive slices.

Compared to electronic radiography, computed tomography provides increased sensitivity and detection capabilities. The contrast resolution of a good quality tomographic image is 0.1 to 0.2%, which is approximately two orders of magnitude better than with x-ray film.

In addition, images are produced in a quantitative ready to use digital format. They provide detailed physical information, such as size, density, and composition, to aid in evaluating defects. Methods are being developed to use this information to predict failure modes or system performance under operating loads. The data can also be easily manipulated to obtain various types of images, to develop automated flaw detection tech-

niques, and to promote efficient archiving. The uses of computed tomography for examining castings are shown in Fig. 11 and 12.

Ultrasonic Inspection

The advantages of ultrasonic tests for castings are:

- High sensitivity, which permits the detection of minute cracks
- Great penetrating power, which allows the examination of extremely thick sections
- Accuracy in measuring of flaw position and estimating defect size

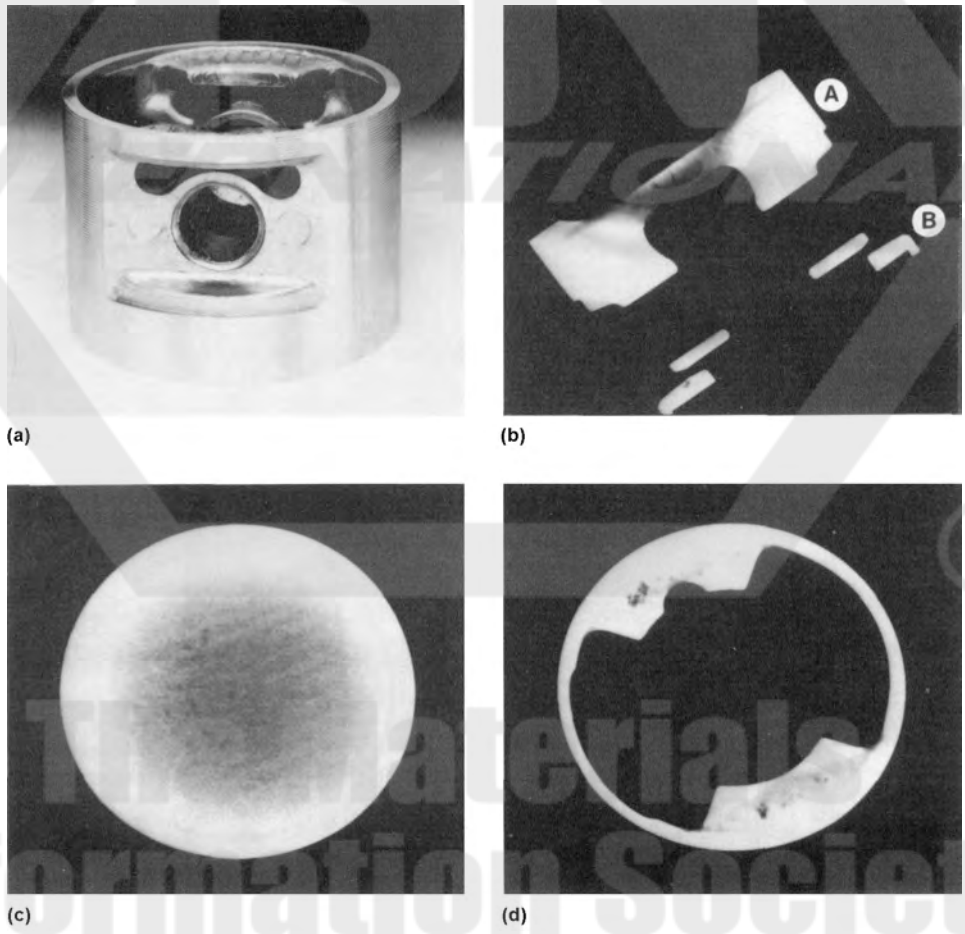


Fig. 11 Computed tomographic images of a die cast aluminum automotive piston. (a) Photograph of cast part. (b) Vertical slice through the piston shows porosity as dark spots in the crown area (point A) and counterbalance area (point B). (c) Transverse slice through the crown of the piston verifies the porosity; the smallest void that is visible is 0.4 mm (0.016 in.) in diameter. (d) Transverse slice through the counterbalance area also verifies porosity. Dimensional analysis of the piston walls is possible to an accuracy of ±50 μm (±0.002 in.). Courtesy of B.G. Isaacson, Bio-Imaging Research, Inc.

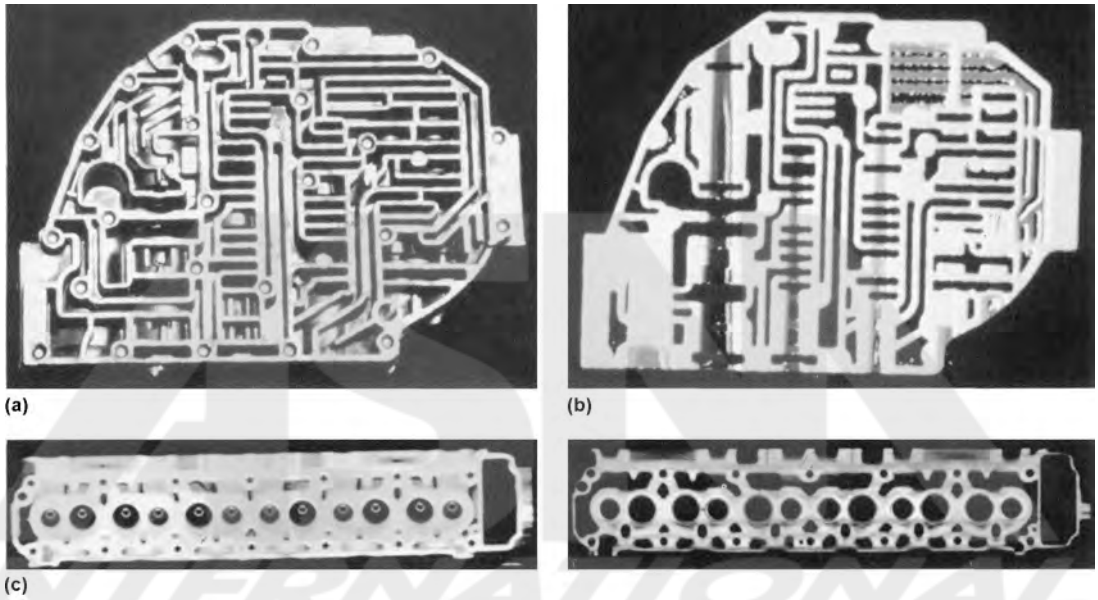


Fig. 12 Use of computed tomography for examining automotive components. (a) Photograph of a cast aluminum transmission case with (b) corresponding tomographic image. (c) Two three-dimensional images of a cast aluminum cylinder head generated from a set of continuous tomographic scans used to view water cooling chambers where a leak had been detected. Courtesy of R.A. Armistead, Advanced Research and Applications Corporation.

Ultrasonic tests have the following limitations:

- Size-contour complexity and unfavorable discontinuity orientation can pose problems in interpreting the echo pattern
- Undesirable internal structure, for example, grain size, structure, porosity, inclusion content, or fine dispersed precipitates, can similarly hinder interpretation
- Reference standards are required

Because castings are rarely simple flat shapes, they are not as easy to inspect as products such as rolled rectangular bars. The reflections of a sound beam from the back surface of a parallel sided casting and a discontinuity are shown schematically in Fig. 13(a), together with the relative heights and positions of the reflections of the two surfaces on an oscilloscope screen. A decrease in the back reflection at the same time as the appearance of a discontinuity echo is a secondary indication of the presence of a discontinuity. However, if the back surface of the casting at a particular location for inspection is not approximately at a right angle to the incident sound beam, the beam will be reflected to remote parts of the casting and not directly returned to the detector. In this case, as shown in Fig. 13(b), there is no back reflection to monitor as a secondary indication.

Many castings contain cored holes and changes in section, and echoes from holes and changes in section can interfere with echoes from discontinuities. As shown in Fig. 13(c), the echo from the cored hole overlaps the echo from the discontinuity on the oscilloscope screen. The same effect is shown in Fig. 13(d), in which echoes from the discontinuity and the casting fillets at a change in section are shown overlapping on the oscilloscope.

Curved surfaces do not permit adequate or easy coupling of the flat search units to the casting surface, especially with contact double search units. This can be overcome to some extent by using suitable viscous couplants, but misleading results may be produced because multiple reflections in the wedge of fluid between the search unit and the surface can result in echoes on the screen in those positions where discontinuity echoes may be expected to appear. Because the reflections inside the couplant use energy that would otherwise pass into the casting, the back echo decreases, and this decrease might be interpreted as confirmation of the presence of a

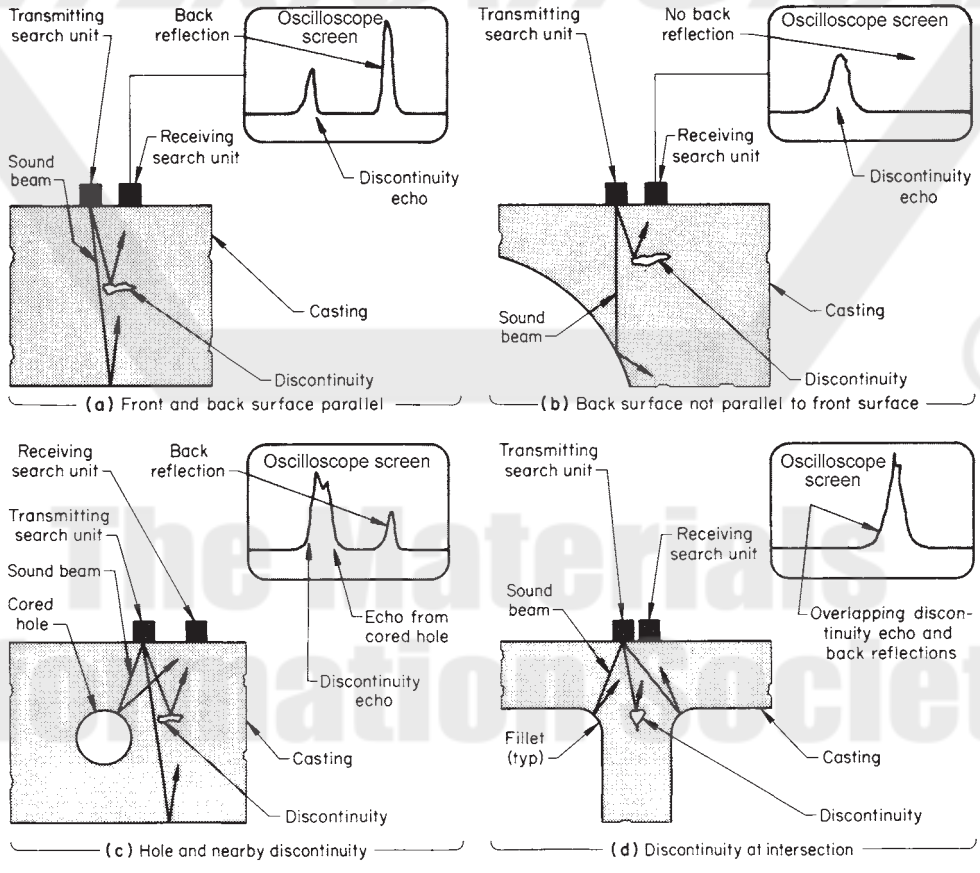


Fig. 13 Schematic of the effect of casting shapes on reflection and oscilloscope screen display of sound beams. See text for discussion. Source: Ref 2

discontinuity. On cylindrical surfaces, the indication will change as a double search unit is rotated. The wedge effect is least when the division between the transmitting and receiving transducers is parallel to the axis of the cylinder. Wedge effects in the couplant are a particular problem on castings curved in two directions. One solution in this case is to use a small search unit so that the wedge is short, although the resolution and sensitivity may be reduced.

If the surface of the casting to be inspected is of regular shape, such as the bore of a cylinder in an engine block, the front of the search unit can be shaped to fit the curvature of the surface. These curved shapes form an acoustic lens that will alter the shape of the sound beam, but unless the curvature is severe, this will not prevent adequate accuracy in the inspection. Cast-on flat metal pads for application of the ultrasonic search unit are very effective and allow particular areas of the casting to be inspected.

Subsurface Defects. Defects, such as small blowholes, pinholes, or inclusions that occur within depths of 3 or 4 mm (0.1 or 0.15 in.) of a cast surface, are among the most difficult to detect. They are beyond the limits of sensitivity of conventional magnetic particle methods and are not easily identified by eddy current techniques. They usually fall within the dead zone (the surface layer that cannot be inspected) of conventional single-crystal ultrasonic probes applied directly to a cast surface, although some improvement can be obtained by using twin crystal probes focused to depths not too far below the surface. The other alternative using contact methods of ultrasonic testing is to employ angle probes, but this complicates the procedures and interpretation methods to the point at which they can only be applied satisfactorily under the close control of skilled operators.

However, freedom from such surface defects is a very important aspect of the casting quality. Apart from their effect in reducing bending fatigue properties, such defects are frequently revealed at late stages in the machining of a component, leading to its rejection. Ultrasonic methods for detecting subsurface defects are much more successful when the dead zone beneath the as-cast surface is virtually eliminated by using immersion methods in which the probe is held away from the cast surface at a known controlled distance, with coupling being obtained through a liquid bath. To make such methods consistent and reliable, the test itself must be automated.

Semiautomatic equipment has been developed for examining castings such as cylinder heads by this method. With this equipment, the casting is loaded into a cradle from a roller track and is then transferred using a hoist into the immersion tank until the surface of the casting to be inspected is just submerged in the liquid. Depth of immersion is closely controlled because the customer will not permit liquid to be left in the internal passageways of the cylinder head. The immersed surface of the casting is

then scanned manually using an ultrasonic probe held at a fixed distance from the casting surface. This equipment is suitable for testing any casting requiring examination over a flat surface.

Internal Defects. Ultrasonic inspection is a well established method for the detection of internal defects in castings. Test equipment developments, automated testing procedures, and improvements in determining the size and position of defects, which is essential to assessing whether or not their presence will likely affect the service performance of the casting, have contributed to the increasing use of ultrasonic test equipment.

For determining the position and size of defects, the usual method of presentation of ultrasonic data is an A-scan, in which the amplitude of the echoes from defects is shown on a time base and has well known limitations. Sizing relies on measuring the drop in amplitude of the echo as the probe is passed over the boundary of a defect or measuring the reduction in the amplitude of the back wall echo due to the scattering of sound by the defect. In most cases, sizing is approximate and is restricted to one or two dimensions. Improvements in data presentation in the form of B-scans and C-scans that present a plane view through the section of the component provide a marked improvement in defining defect positions and size in two or three dimensions. Such displays have been used for automated defect characterization systems in which porosity, cracks, and dross have been distinguished. Because of the requirement to scan the probe over the surface, the application of B-scan and C-scan methods has generally been limited to simple geometric shapes having good surface finish, such as welded plate structures. Application to castings is currently restricted, but greater use of B-scan and C-scan methods is likely with either improved scanning systems or arrays of ultrasonic probes.

Structure evaluation is an area of growing importance for foundry engineers. Ultrasonic velocity measurements are widely used as a means of guaranteeing the nodularity of the graphite structure, and if the matrix structure is known to be consistent, guaranteeing the principal material properties of ductile irons. Velocity measurements have been used to evaluate compacted graphite iron structures to ensure that the desired properties have been consistently obtained.

Leak Testing

Castings that are intended to withstand pressures can be leak tested at the foundry. Various methods are used, according to the type of metal being tested. One method consists of pumping air at a specific pressure into the inside of the casting in water at a given temperature. Any leaks through the casting become apparent by the release of bubbles of air through the faulty portions. An alternative method is to fill the cavities of a casting with paraffin at a specified pressure. Paraffin, which penetrates the smallest of crevices, will rapidly find any defect, such as porosity, and

will show quickly as an oily or moist patch at the position of the fault. Liquid penetrants can be poured into areas of apparent porosity and time allowed for the liquid to seep through the casting wall. However, the introduction of contaminants into the defect may make repair welding more difficult.

The pressure testing of rough (unmachined) castings at the foundry may not reveal any leaks, but it must be recognized that subsequent machining operations on the casting may cut into porous areas and cause the casting to leak after machining. Minor seepage leaks can be sealed by impregnation of the casting with liquid or by filling with sodium silicate, a synthetic resin, or other suitable substance. As-cast parts can be impregnated at the foundry to seal leaks if there is little machining or if experience has shown that machining does not affect the pressure tightness. However, it is usually preferable to impregnate the casting after final machining.

ACKNOWLEDGMENT

This chapter was adapted from Inspection of Castings, *Nondestructive Evaluation and Quality Control*, Volume 17, *ASM Handbook*, 1992.

REFERENCES

1. R. DasGupta, Common Defects in Various Casting Processes, *ASM Handbook*, Vol 15, *Casting*, ASM International, 2008, p 1192–1202
2. Inspection of Castings, *Nondestructive Evaluation and Quality Control*, Vol 17, *ASM Handbook*, ASM International, 1989, p 231–277

CHAPTER 13

Inspection of Steel Bar and Wire

THIS CHAPTER will focus on the inspection of steel bars; however, the principles involved also apply, for the most part, to steel wire. In many cases, as far as nondestructive inspection is concerned, steel bars and wire are the same.

The primary objective during the inspection of steel bars and wire is generally the same as for the inspection of other products, which is to detect conditions in the material that may be detrimental to the satisfactory end use of the product. However, there is an additional objective in attempting to detect undesirable conditions in semifinished products, namely, to eliminate unacceptable material before spending time, money, and energy in manufacturing products that will later be rejected.

The inspection of bars and other semifinished products does not impair the product, provides rapid feedback of information, and can be utilized as either an in-line or off-line system. It makes use of several devices, such as visual, audio, and electromagnetic, for the detection of flaws and of variations in composition, hardness, and grain structure. A wide range of selectivity is provided for each device, permitting acceptance or rejection at various specification levels. The most common function of inspection in the steel industry is the detection and evaluation of flaws. It is also used for the detection of variations in composition and physical properties. However, no amount of inspection can ensure an absolutely flawless bar, but it does provide a consistent specified degree of quality during everyday operation.

Types of Flaws Encountered

The terms used for the various types of flaws discussed may not be the same in different geographic areas. In many cases, different terms are ap-

plied to the same type of flaw. Therefore, this section contains a description and an illustration of each condition. The term *flaw* is applied to blemishes, imperfections, faults, or other conditions that may nullify acceptability of the material. The term also encompasses such terms as *pipe*, *porosity*, *laminations*, *slivers*, *scabs*, *pits*, *embedded scale*, *cracks*, *seams*, *laps*, and *chevrons*, as well as *blisters* and *slag inclusions* in hot rolled products. For products that are cold drawn, *die scratches* may be added.

Most flaws in steel bars can be traced back to the pouring of the hot metal into molds. Factors that work against obtaining a perfect homogeneous product are:

- The fast shrinkage of steel as it cools (roughly 5% in volume)
- The gaseous products that are trapped by the solidifying metal as they try to escape from the liquid and semisolid metal
- Small crevices in the mold walls, which cause the metal to tear during the stripping operation
- Spatter during pouring, which produces globs of metal frozen on the mold walls because of the great difference in temperature of the mold surfaces and the liquid metal

Pipe is a condition that develops in the nominal top centerline of the ingot as the result of solidification of the molten metal from the top down and from the mold walls to the center of the ingot (Fig. 1). Because of the metal shrinkage and lack of available liquid metal, a cavity develops from the top down and, if not completely cropped before subsequent rolling, becomes elongated and will be found in the center of the final product, as shown in ingot B in Fig. 1.

Porosity is the result of trapped gaseous bubbles in the solidifying metal causing porous structures in the interior of the ingot (Fig. 1). On rolling, these structures are elongated and interspersed throughout the cross section of the bar product.

Inclusions may be the products of deoxidation in the ingot, or they may occur from additives for improving machinability, such as lead or sulfur. Inclusions and their typical location in a steel bar are shown in Fig. 2(a).

Laminations may occur from spatter (entrapped splashes) during the pouring of the steel into the mold. They are elongated during rolling and are usually in the bar subsurface. A lamellar structure opened up by a chipping tool is illustrated in Fig. 2(b).

Slivers are most often caused by a rough mold surface, overheating prior to rolling, or abrasion during rolling. Very often, slivers are found with seams. Slivers usually have raised edges, as shown in Fig. 2(c).

Scabs are caused by splashing liquid metal in the mold. The metal first freezes to the mold wall, then becomes attached to the ingot, and finally becomes embedded in the surface of the rolled bar (Fig. 2d). Thus, scabs bear some similarity to laminations.

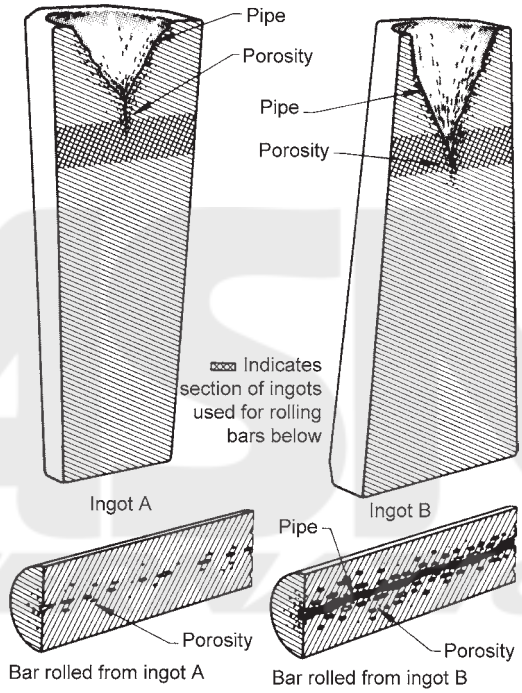


Fig. 1 Longitudinal sections of two types of ingots showing typical pipe and porosity. When the ingots are rolled into bars, these flaws become elongated throughout the center of the bars. Source: Ref 1

Pits and Blisters. During subsequent rolling, gaseous pockets in the ingot often become pits or blisters on the surface or slightly below the surface. Other pits may be caused by overpickling to remove scale or rust. Pits and blisters are both illustrated in Fig. 2(e).

Embedded scale may result from the rolling or drawing of bars that have become excessively scaled during prior heating operations. The pattern illustrated in Fig. 2(f) is typical.

Cracks and seams are often confused with each other. Cracks with little or no oxide present on their edges may occur when the metal cools in the mold, setting up highly stressed areas. Seams develop from these cracks during rolling as the reheated outer skin of the billet becomes heavily oxidized, transforms into scale, and flakes off the part during further rolling operations. Cracks also result from highly stressed planes in cold drawn bars or from improper quenching during heat treatment. Cracks created from these latter two causes show no evidence of oxidized surfaces. A typical crack in a bar is shown in Fig. 2(g).

Seams result from elongated trapped gas pockets or from cracks. The surfaces are generally heavily oxidized and decarburized. Depth varies widely, and surface areas sometimes may be welded together in spots. Seams may be continuous or intermittent, as indicated in Fig. 2(h). A micrograph of a typical seam is shown in Fig. 3.

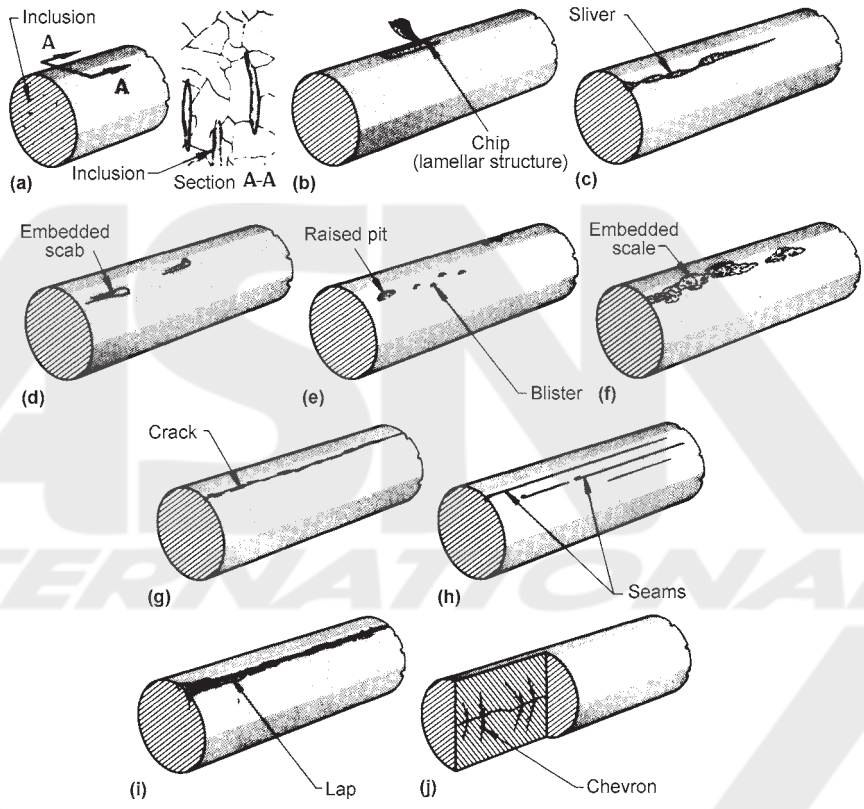


Fig. 2 Ten different types of flaws that may be found in rolled bars. See text for discussion. Source: Ref 1

Laps are most often caused by excessive material in a given hot roll pass being squeezed out into the area of the roll collar (Fig. 2(i)). When turned for the following pass, the material is rolled back into the bar and appears as a lap on the surface.

Chevrons are internal flaws named for their shape (Fig. 2(j)). They often result from excessively severe cold drawing and are even more likely to occur during extrusion operations. The severe stresses that build up internally cause transverse subsurface cracks.

Methods Used for Inspection of Steel Bars

In addition to a thorough visual inspection of steel bars, four methods used either singly or in combination are:

- Magnetic particle inspection
- Liquid penetrant inspection
- Ultrasonic inspection
- Electromagnetic inspection

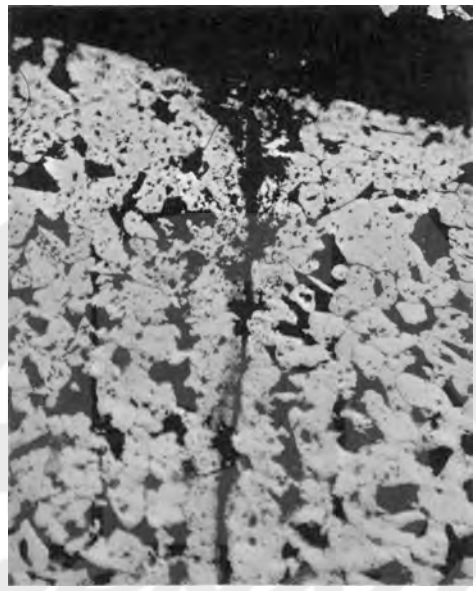


Fig. 3 Micrograph of a seam in a cross-section of a 19 mm (¾ in.) diameter medium carbon steel bar showing oxide and decarburization in the seam. 350×. Source: Ref 1

Magnetic Particle Inspection

Magnetic particle inspection offers the same visual aid in the nondestructive inspection of bars as it does for castings, forgings, or machined products. The method is used for detecting seams, cracks, and other surface flaws, and, to a limited extent, subsurface flaws. As a rule, the method is not capable of detecting flaws that are more than 2.5 mm (0.1 in.) beneath the surface.

Longitudinal Flaws. Optimum indications are obtained when the magnetic field is perpendicular to the flaws. A similar result is obtained for flaws slightly below the surface, but the surface leakage is less and, consequently, fewer iron particles are attracted to the area, producing a less definite indication.

Various colors of iron powders are commercially available to permit the choice of a color that provides maximum contrast between the powder and the material being inspected. Fluorescent coatings on powders and ultraviolet light can be used to make the indication more vivid. The powders can be applied in dry form, or can be suspended in oil or a distillate and flowed over the workpiece during or after the magnetizing cycle.

Transverse Flaws. To detect flaws transverse to the long axis of the bar being inspected, a solenoid winding or encircling coil is used. To protect the bar from arc burns when the current is turned on, electrical contact is usually made by soft metallic pads held firmly against the bar ends.

The power used can be either direct current or alternating current. Direct current may be from batteries or rectified alternating current. Alternating current travels near the surface and should not be used for detecting subsurface flaws. In most cases, the continuous magnetization system is used for bars because most bars have low retentivity for magnetism; therefore, the residual magnetism system is not suitable. Finished bars must be demagnetized; otherwise, during manufacturing operations such as machining, steel chips will adhere and possibly cause trouble.

As a rule, the magnetic particle inspection of bars is confined to the inspection of a small quantity of bars, as in a fabricating shop. The method is considered too slow and too costly for mass production inspection, as at the mill.

Liquid Penetrant Inspection

Liquid penetrant inspection (another visual aid), for several practical reasons, is not extensively used for detecting flaws in steel bars. These reasons include the following:

- Its use is restricted to the detection of flaws that are open to the bar surface
- Adaptation to automation is limited compared to certain other inspection methods
- Time cycles are too long for the inspection of bars on a mass production basis

However, there are exceptions and there are cases where liquid penetrant inspection has been used for inspecting from one to a few bars, as in a fabricating shop. Specific advantages are:

- Liquid penetrant inspection is extremely sensitive and can sometimes detect surface flaws missed by other methods
- The solvent-removable system (one of the several liquid penetrant systems) in particular is extremely flexible and can be used for inspecting bars or portions of bars in virtually any location, including in the field

Ultrasonic Inspection

Ultrasonic inspection is done with high-frequency (about 1 to 25 MHz) sound waves and can successfully detect internal flaws in steel bars. Most often, the ultrasonic inspection of steel bars is restricted to large diameter bars and to applications where high integrity is specified. Also, because of the limitations of ultrasonic inspection for detecting surface flaws, it is ordinarily used in conjunction with some other method that is more suitable for inspecting bar surfaces.

An ultrasonic beam has the valuable property that it will travel for long distances practically unaltered in a homogeneous liquid or solid, but when

it reaches an interface with air (for example, at a crack or at the surface of a metal body), it is almost completely reflected.

The technique most commonly used for the inspection of bars or bar-like workpieces is the pulse echo technique. Short pulses of ultrasonic energy are passed through the bar. The sweep voltage of the time base is coordinated with the pulse repetition frequency so that the reflections are indicated on an oscilloscope screen. A certain amount of energy is reflected at the interface between the probe and workpiece, giving the first transmission signal. The probe can either have two separate crystals, a transmitter and a receiver, or have only one, which is used alternately as transmitter and receiver.

The ultrasonic method is characterized by high sensitivity and very deep penetration, but in addition to its surface limitations, its production speed is relatively low. A liquid couplant is necessary and can be a source of interference. This method is suitable for testing ingots, billets, plate, and tubes in addition to bars or bar-like workpieces. In certain cases, ultrasonic inspection has been automated. Typical products that are ultrasonically inspected using automated equipment are forged axle shafts (which are, in effect, extruded bars) and rolled bars.

Cold Drawn Bars. The most effective method for the inside flaw inspection of cold drawn bars is ultrasonic flaw detection. However, it is necessary to detect the smaller defects in the near surface area. The conventional normal beam method (Fig. 4a) is not satisfactory, because of the untested area near the surface.

A testing method for detecting smaller flaws immediately under the surface of cold drawn bars is the angle beam method (Fig. 4b), which conveys ultrasonic waves into the material with an angle beam. It can detect the flaws immediately under the surface that are in the dead zone for the

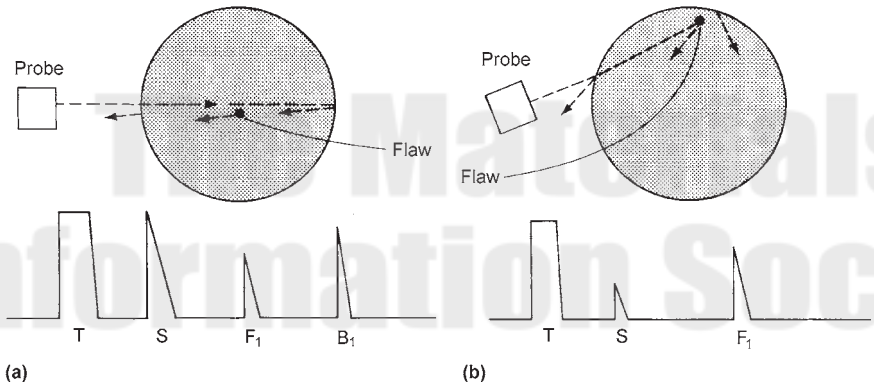


Fig. 4 Schematic showing position of probe relative to flaw inside of bar and resulting wave display obtained for two methods of ultrasonic flaw detection. (a) Normal beam method. (b) Angle beam method. Wave display nomenclature: T, transmit pulse; S, surface reflection echo; F_1, flaw echo; B_1, back wall echo. Source: Ref 2

conventional normal beam method. Entire cross-sectional area testing becomes possible with the angle beam method and the conventional normal beam method in combination. The testing method to feed the material spirally and to make the probes follow the deflection of the material feeding has already been adopted in practical use for as-rolled steel bars. It is difficult to obtain higher testing speed for cold drawn bars because of smaller dimensions. Therefore, a method has been developed in which the material is fed straight and the probes are simultaneously rotated at high speed. Table 1 lists the main system specifications, and a schematic of the setup is shown in Fig. 5. For bars with smaller dimensions, guide sleeves and tripplet rollers are used to prevent the ultrasonic incident angle to the material from changing because of excessive vibration and/or bending of the material. The water circulation system also incorporates a device that stabilizes the coupling water.

For flaws located immediately under the surface, the angle beam method record can detect flaws as small as 0.2 to 0.3 mm (0.008 to 0.012 in.). Flaw echoes this small are not detectable with the normal beam method.

Table 1 Specifications of a rotating type ultrasonic flaw detection system

Parameter	Specifications
Dimension of material, mm (in.)	15–32 (0.59–1.26)
Testing method	Normal-beam method and angle-beam method
Testing frequency, MHz	10 and 5
Number of rotations of probe, rev/min	1000
Signal transmit	Noncontact rotation transmit
Marker	One each for near-surface flaw and inside flaw

Source: Ref 2

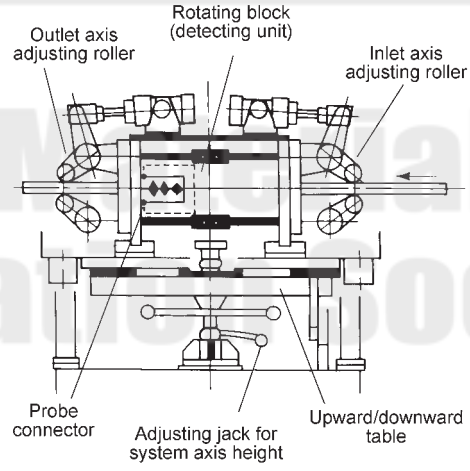


Fig. 5 Schematic of a typical rotating type ultrasonic flaw detection system. Source: Ref 2

Cold Drawn Hexagonal Bars. Requirements for strict quality assurance are required for gaging inside flaws to the same level as surface flaws. The conventional testing method is manual detection with the normal beam method. Because this method requires testing with plural directions, working efficiency is low. Furthermore, an untested zone remains at the area immediately under the surface. Therefore, a testing system using the entire cross-section with higher efficiency has been sought.

Higher efficiency has been attained by incorporating an automated ultrasonic flaw detection system with probes for each face of the material to detect separately the flaws located on the inside area and the near-surface area (Fig. 6). Flaws inside the material are detected with the normal beam method at each face of the material. In this method, the untested zone remains in the near-surface area. Therefore, surface and near-surface area flaws are detected with the angle beam method at each face of the material. That is, six normal beam probes and six angle beam probes are located on the circumference of the materials to be tested, which is conveyed longitudinally. The probe positions are arranged so that the entire cross-section can be detected.

The probe holder is designed so that all the probes can be adjusted simultaneously by adjusting one when the material size is changed. The coupling medium is a special oil that has low ultrasonic attenuation and causes no rust on the material to be tested. The specifications of the system are listed in Table 2. Flaws larger than 0.3 mm (0.012 in.) can be detected at the near-surface area. Flaws measuring at least 0.2 mm (0.008 in.) can be detected deep inside the hexagonal bar material.

Ultrasonic Flaw Detection on Cold Drawn Wires. Surface flaw inspection is important for drawn wires. A rotation type eddy current flaw

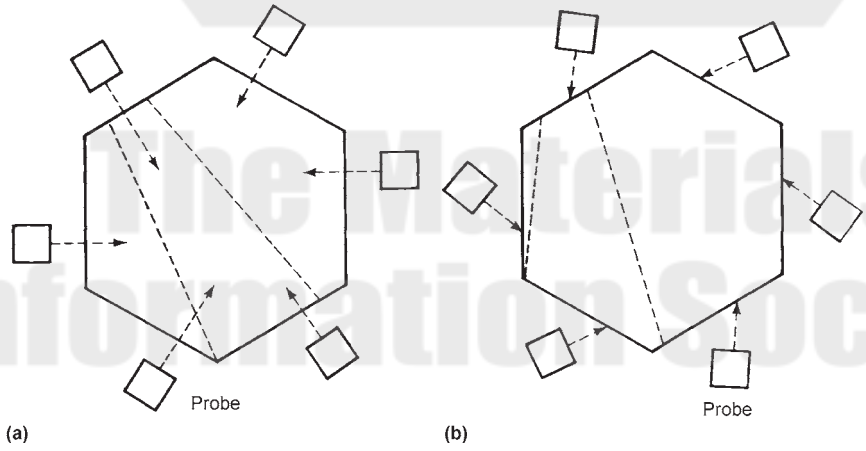


Fig. 6 Dual set of six circumferentially mounted probes used to ultrasonically detect flaws in cold drawn hexagonal bars. (a) Normal beam method to detect flaws deep inside bar. (b) Angle beam method to detect surface and near-surface flaws. Source: Ref 2

detection system is used for quality assurance. For drawn wires, rotating type eddy current flaw detection has been used in combination with rotating ultrasonic flaw detection to detect surface defects and inside flaws, respectively, in a two step process. However, the high cost and inefficiency of this method have prompted the development of a system with a rotating type ultrasonic flaw detection unit that can also detect surface flaws. An additional die is placed behind the cold drawing die to stabilize the vibration of the material. A detection unit, which has probes arrayed in a circumferential direction, is placed between these dies. The three detection modes (Fig. 7) are:

- Surface wave detection mode for surface defects
- Angle beam detection mode for near-surface defects
- Normal beam detection mode for inside defects

The ultrasonic incident angle can be optimized according to material dimensions. Water, the coupling medium, is always kept in full quantity even in high speed rotation. Thus, the system has the stable mechanism to provide constant detection.

Table 2 Specifications of an ultrasonic flaw detection system for cold drawn hexagonal bars

Parameter	Specifications
Dimension of material, mm (in.)	12–32 (0.472–1.260)
Testing method	Normal-beam, 6 channels; angle-beam, 6 channels
Testing frequency, MHz	5
Probe position	Fixed in circumferential direction
Marker	Two for near-surface flaw and inside flaw

Source: Ref 2

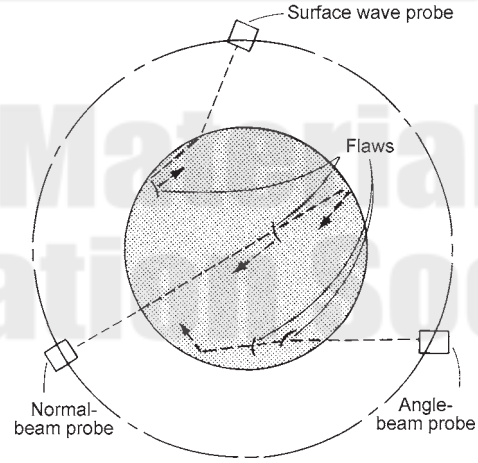


Fig. 7 Principle of ultrasonic flaw detection for cold drawn wires using three detection mode probes. Source: Ref 2

One advantage of this system is that linear defects can be detected by surface wave detection at the same level as an eddy current method. Another is that the entire cross section can be covered by means of a combination angle beam/normal beam method. The specifications of this setup are summarized in Table 3. Results with this system showed detectability of 0.1 mm (0.004 in.) minimum flaw depth on surface defects and 0.2 mm (0.008 in.) minimum inside defect size. This system enables the user to inspect the entire cross section of cold drawn wires to a high degree of accuracy.

Electromagnetic Inspection Methods

Electromagnetic methods of inspection are used far more extensively for nondestructive inspection of steel bars than any of the methods discussed. Electromagnetic methods are readily adaptable to automation and can be set up to detect flaws, as well as a number of different compositional and structural variations, in bars on a mass production basis.

Equipment can be relatively simple, but for mass production inspection, the equipment may be highly sophisticated and costly. Such equipment can not only detect flaws and indicate them on an oscilloscope or other form of readout, but can also mark the location of the flaw on the bar before it emerges from the inspection equipment and can automatically sort the bars on the basis of seam depth.

Electromagnetic systems include the systems that use magnetic fields generated by alternating current flowing in a solenoid. A wide range of frequencies is used. As the alternating current flows through the solenoid, the magnetic field generated induces eddy currents within the metal workpiece. These currents are affected by the electrical resistivity (more commonly expressed as electrical conductivity, the reciprocal of resistivity), magnetic permeability, configuration, homogeneity, surface irregularities, and flaws in the metal. The resistivity of the workpiece can vary because of the chemical composition, crystal orientation, structure, and history of mechanical working. Permeability will vary over a broad range, depending on the amount of stresses present in the workpiece. It increases slightly in the vicinity of a flaw when the bar is subjected to a stress producing operation.

Table 3 Specifications of an ultrasonic flaw detection system for cold drawn wires

Parameter	Specifications
Dimension of material, mm (in.)	15–30 (0.590–1.181)
Testing frequency	Normal beam: 10 MHz, 1 channel Angle beam: 5 MHz, 2 channels Surface wave: 5 MHz, 2 channels
Number of rotations of probe, rev/min	1000
Signal transmit	Noncontact rotation transmit
Marker	One each for near-surface flaw and inside flaw

Source: Ref 2

Electromagnetic systems of flaw detection are broadly classified as:

- Those depending primarily on variations in electrical conductivity
- Those depending primarily on variations in magnetic permeability

Both systems are capable of detecting flaws in ferromagnetic bars. The conductivity dependent systems can also be used to detect flaws in nonferromagnetic bars.

Eddy Current Systems

When electrical conductivity (resistivity) is the major variable, the test procedure is known as the eddy current system. The alternating field intensity is low, permitting the use of a correspondingly small inductor. Most eddy current systems use a constant voltage input derived from an electronic oscillator with a means of varying the output frequency through a wide range, such as from 0.5 to 1000 kHz, in discrete steps.

For general flaw detection, the range of 1 to 50 kHz is widely used. For ferromagnetic bars, a means must be provided to eliminate or minimize the effects of permeability variation. This is usually accomplished by magnetically saturating the bar being tested. The means for doing this is either a dc solenoid or a strong permanent magnet. A longitudinal section of one type of eddy current coil assembly is shown in Fig. 8, and a more detailed drawing of the rotating coil setup is shown in Table 4.

Eddy current inspection is especially useful for detecting and evaluating seams in steel bars. With this system, depending on the circuitry used, a difference of as little as 0.025 mm (0.001 in.) in seam depth can be detected. Because of the skin effect, the ability of eddy currents to penetrate the test metal decreases in proportion to the increases of the frequency.

Eddy current inspection can be used without magnetic saturation for inspecting hot bars in the mill when the metal is above the Curie tempera-

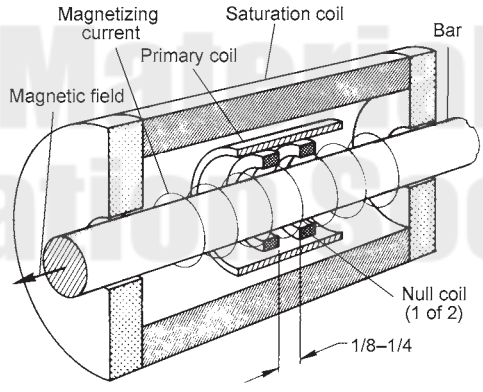
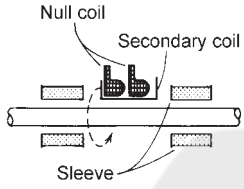


Fig. 8 Coil assembly for the inspection of steel bars by the eddy current system. Dimensions in inches. Source: Ref 1

Table 4 Specifications of a rotating probe type eddy current flaw detection system



Parameter	Type I	Type II
Dimension of material, mm (in.)	5–32 (0.197–1.260)	5–25 (0.197–0.984)
Number of probes	2	4
Probe area, mm^2 ($in.^2$)	10 (0.016)	5 (0.0078)
Number of rotations of probe, rev/min	3000	6000
Testing frequency, kHz	64	512
Signal transmit	Noncontact rotation transmit	Noncontact rotation transmit

Source: Ref 2

ture, because the metal is nonmagnetic at this temperature. Therefore, it follows that the magnetic permeability system cannot be used to inspect hot bars.

Eddy Current Testing of Cold Drawn Bars. Surface defects on cold drawn bars can be inspected by eddy current detection methods using an encircling coil. This method utilizes a rotating probe that detects surface defects with the probe coil rotating at high speed around the circumference of the cold drawn bars.

The encircling coil method exhibits lower detectability on linear flaws because flaw detection depends on the difference between two test coils in which the material to be tested is encircled. On the other hand, the method of rotating the probe coil at high speed along the circumference of the material can detect linear defects because it detects bars in spiral scanning.

The specifications of the detection system are listed in Table 4. One of the main features is signal transmission in the probe rotation unit by the noncontact rotating transmit method, which requires no maintenance work. Guide sleeves are placed in front of and behind the probe to maintain a constant distance between the probe and the material to be tested, which is important for acceptable performance (Fig. 9). Furthermore, the rotation axis of the probe and the axis of the workpiece are kept in a line by pinch rollers placed in front of and behind the detector. On the probe, a distance sensor is used for the automatic gain control function to provide electric compensation against distance variation.

The relation between flaw depth and signal output is shown in Fig. 10. Natural flaws produce a larger deviation in signal output than artificially introduced flaws because of the complicated cross-sectional configuration of the flaw, but the minimum detectable flaw depth is 0.1 mm (0.004 in.). Detectable flaw length depends on the feeding speed of the material, the number of probes, and the number of rotations. For example, at a speed of

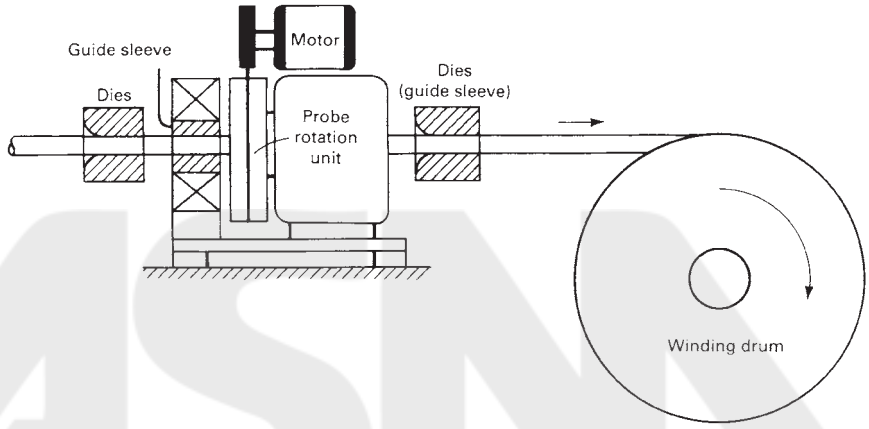


Fig. 9 Schematic of a rotating probe type eddy current flaw detector. Source: Ref 2

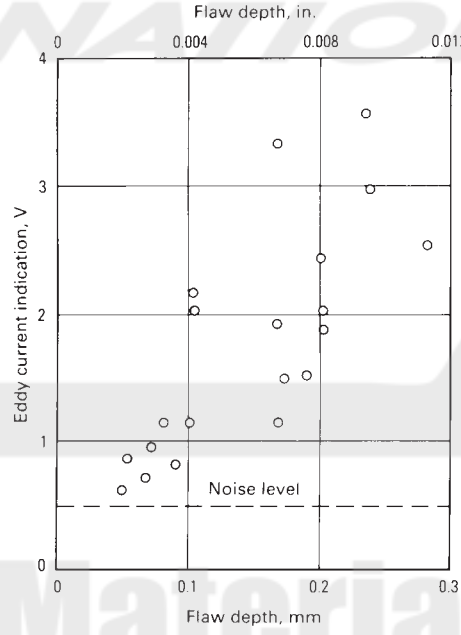


Fig. 10 Plot of eddy current signal output versus flaw depth to gage detectability of flaws in cold drawn bars. Source: Ref 2

60 m/min (200 sfm), the full surface is converted, and the minimum detectable flaw length is as long as the length of the probe coil.

Eddy Current Flaw Detection on Cold Drawn Hexagonal Bars. Cold finished steel profiles (hexagonal bars) are mainly used as the raw material for couplers in oil pressure piping, an application for which quality assurance is important. Surface defects on cold drawn hexagonal bars include cracks derived from the cold working process as well as material

flaws, both of which are long, longitudinal defects. It is impossible to detect these defects by the differential method using encircling coils. The rotating probe method is also not applicable, because of the hexagonal form. An automated flaw detection system for cracks initiated by the working process was developed using the eddy current flaw detection system by a standard voltage comparison method.

There are two methods for testing cold finished steel hexagonal bars: the standard voltage comparison method with encircling probes (Fig. 11b) and the differential method with probe assembly (Fig. 11c). There is no effective difference in detectability between these two methods. For the probe assembly method, it is necessary to consider the differences in detectability of each individual probe, which is not necessary for the standard voltage comparison method.

The standard voltage comparison method is inferior in detectability to the rotating probe method, but is less expensive and can efficiently detect cracks resulting from the cold working process. This method, which can detect material flaws more than 0.6 mm (0.024 in.) deep, is illustrated in Fig. 12.

Eddy Current Flaw Detection of Cold Drawn Wires. Surface flaw detection on wire drawing line has been conducted by the encircling type

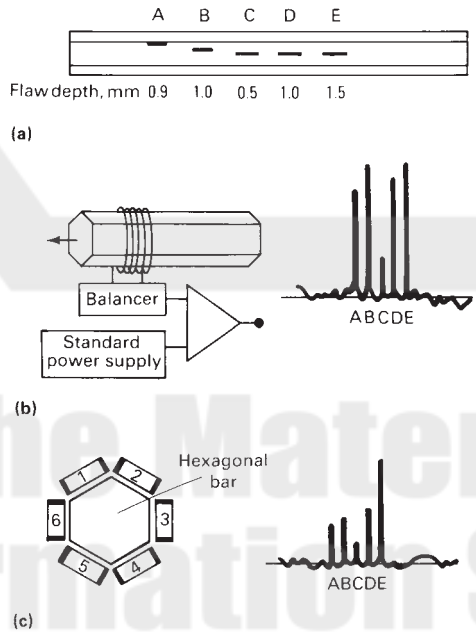


Fig. 11 Eddy current flaw detection method for cold-drawn hexagonal bars. (a) Location of artificial flaws ranging from 0.5 to 19 mm (0.020 to ¾ in.) below probe position. (b) Schematic of setup for standard voltage comparison (encircling coil) method (left) and plot of signals obtained for the designated flaw depths (right). (c) Schematic of setup for differential (six probe coil assembly) method (left) and plot of signals obtained for the designated flaw depths (right). Source: Ref 2

eddy current method. However, this method has difficulty in detecting linear flaws. A rotating probe type eddy current detection method can be effective, as illustrated in Fig. 9 for use on cold drawn bars. It is important in the rotating probe method to maintain a constant distance between the probe and the material to be tested. The rotating unit is positioned between dies where the smaller vibration of the material is expected. Guide sleeves are used to adjust the rotating axis and the axis of the material to be tested. Detectability is illustrated in Fig. 13. Flaws having a 0.1 mm (0.004 in.) minimum depth are detectable.

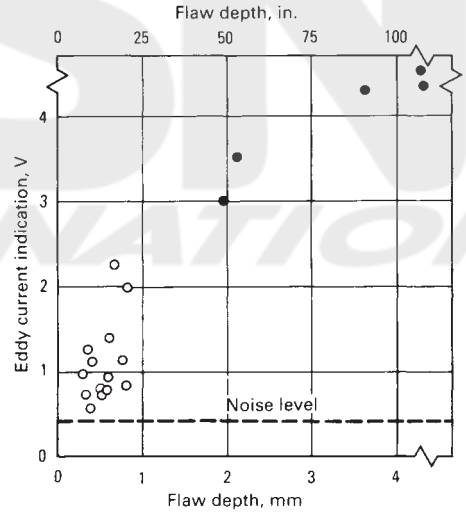


Fig. 12 Plot of eddy current signal output versus flaw depth to measure detectability of flaw, specifically material flaws (open circles) and process induced cracks (closed circles), in cold drawn hexagonal bars. Source: Ref 2

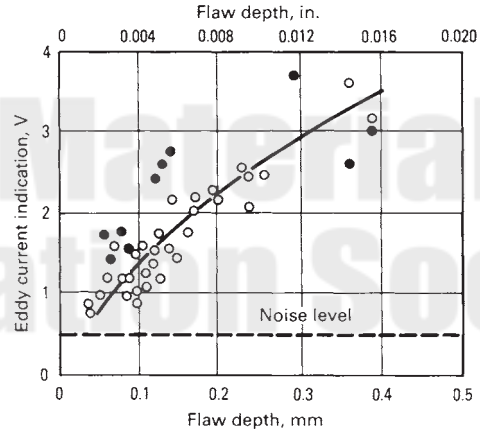


Fig. 13 Plot of eddy current signal output versus flaw depth to measure detectability of flaws, specifically cracks (open circles) and scabs (closed circles), in cold drawn wires. Source: Ref 2

Eddy Current Flaw Detection for a Cold Forged, High Tensile Sheared Bolt. A general view of a high tension sheared bolt is shown in Fig. 14. This type of bolt has a head with a round cross section and is mainly used for general construction and bridge applications. This bolt is produced by cold forging from cold drawn wires in the diameter similar to the outside diameter of a threaded part of the bolt. The head is the most severely processed part of the bolt. The circumferential part of the bolt head is formed between punch and die during cold forging; therefore, cracks tend to occur on the head. Eddy current testing can detect flaws in the bolt head at high speed with the probe rotating method.

A general view of the inspection system used is shown in Fig. 15, and the main specifications are listed in Table 5. Bolts are conveyed from hopper to the line-up unit. Lined up bolts are conveyed to the index table by straight feeder and then conveyed intermittently to the rotating detection head and further to the separator.

After the bolt heads are detected with the rotating detection head, the bolts are classified as good/no-good and separated according to detection

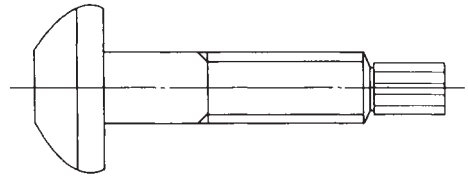


Fig. 14 Schematic of a high tension sheared bolt. Source: Ref 1

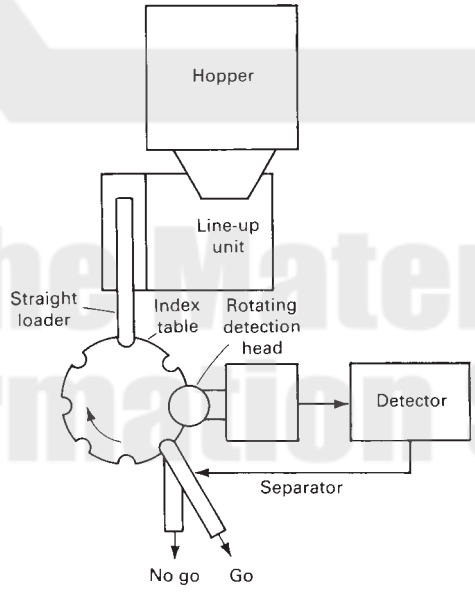


Fig. 15 Schematic of eddy current flaw detection system used to inspect sheared bolt illustrated in Fig. 14. Source: Ref 2

result. The operation of the rotating detection head is shown in Fig. 16. The rotating detection head repeats the following operations while rotating continuously regardless of the position of the bolt head to be tested:

- A bolt stops immediately under the detection head (Fig. 16a)
- The detection head descends while maintaining rotation (Fig. 16b)
- The detection head approaches the bolt head, scans around the bolt head for two revolutions, and detects any flaws (Fig. 16c)
- Pincer-like probe holders release from the bolt head, and the detection head ascends
- Bolt is conveyed to separator while next bolt is conveyed to the position immediately under detection

A detection rate of 60 pieces per minute was maintained by the mechanism to keep the detection head rotating continuously. The relation be-

Table 5 Specifications of an eddy current detection system for a high tension sheared bolt

Parameter	Specification
Material	M20
Testing speed, pieces/min	60
Number of rotations of detecting head, rev/min	300
Testing frequency, kHz	125
Probe type	Self-induction, self-comparison

Source: Ref 2

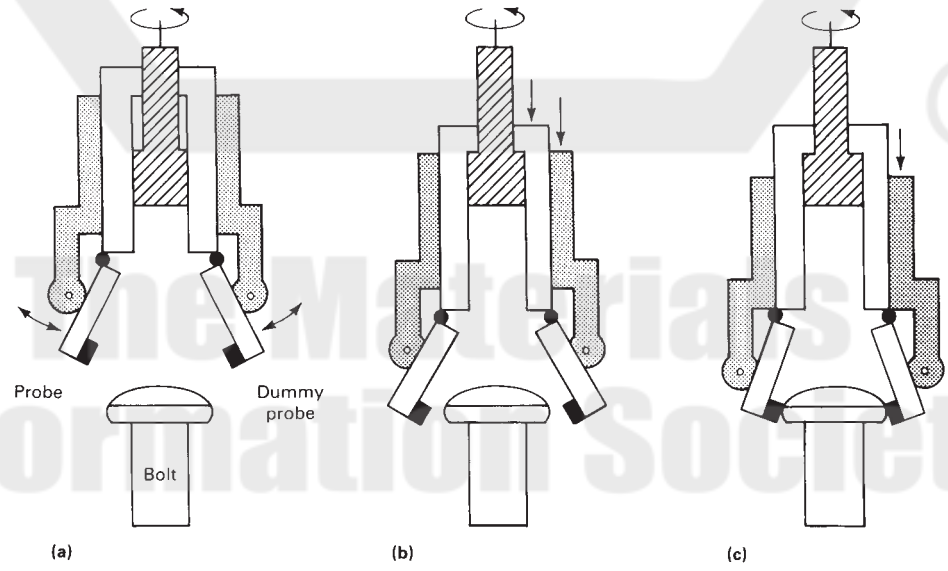


Fig. 16 Operation of rotating eddy current detection head. (a) Shear bolt positioned under rotating detection head. (b) Rotating detection head descends to lower probes into position to inspect bolt head. (c) Probe scans bolt head as bolt undergoes two complete revolutions to detect flaws. Source: Ref 2

tween flaw depth and signal output is shown in Fig. 17. Noise level is high at the circumferential surface of the bolt head because of surface roughness, but the minimum detectable flaw depth is 0.3 mm (0.012 in.).

Magnetic Permeability Systems

Systems that depend on variations in magnetic permeability can be used for detecting flaws and for detecting differences in composition, hardness, or structure. With appropriate instrumentation, both functions can be accomplished simultaneously.

Magnetic permeability systems usually employ a solenoid (primary coil), which is excited by the standard line frequency of 60 Hz with an adjustable current control to produce magnetic fields from 1000 to 30,000 ampere-turns; however, the solenoid is usually operated in the range of 10,000 to 15,000 ampere-turns. A typical coil arrangement used for permeability systems is shown in Fig. 18.

As shown in Fig. 18, the coil arrangement consists of a primary coil (60 Hz), two null coils (zero voltage output coils), and two standard coils. The secondary or pickup coils (null coils) are concentric with the primary coil, connected electrically in opposition, and adjusted to a null or zero voltage output. The null coils are usually spaced 75 to 102 mm (3 to 4 in.) apart. The reason for this spacing is that a normal seam in a bar tapers into the bar to sound material. The variation in stress level producing a measurable change in magnetic permeability is related to the change in seam depth found usually within 75 mm (3 in.) of seam length.

The detection of flaws by permeability systems depends on permeability variations resulting from changes in stress, due to cold work or heat

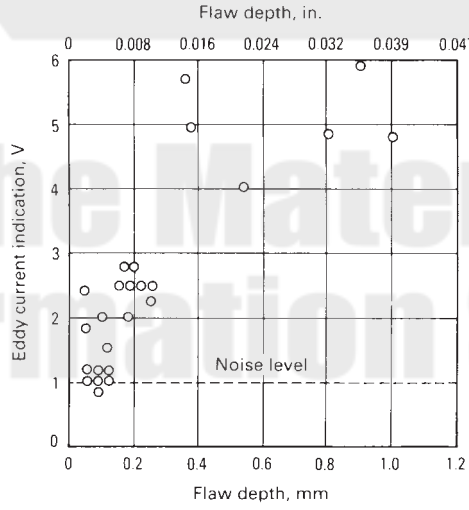


Fig. 17 Plot of eddy current signal output versus flaw depth to measure detectability of flaws in high tensile sheared bolts. Source: Ref 2

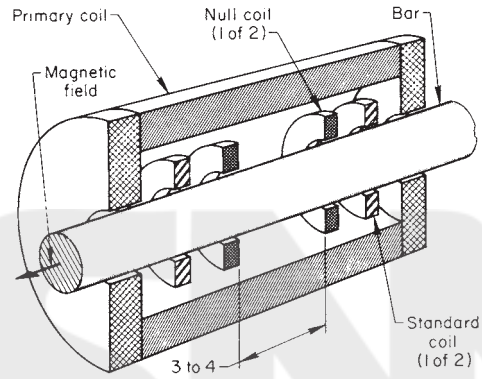


Fig. 18 Coil assembly used for the simultaneous detection of flaws and of variation in composition, structure, and hardness in steel bars. Dimensions in inches. Source: Ref 1

treatment, in the adjacent area of the flaw. These changes are more or less directly proportional to the change in stress up to the elastic limit of the ferrous product.

These systems cannot be used to inspect hot rolled or annealed bars unless they have been subjected to some uniform cold work, such as rotary straightening for round material or planar type straightening for square, hexagonal, or flat sections. Gage straightened bars are not suited to inspection by permeability systems, because of nonuniform high stress concentrations wherever the ram meets the work metal. Such stresses are far in excess of those for flaws in uniformly stressed material.

The efficiency of flaw detection is a function of uniform residual stress levels within the bar. The five conditions in order of decreasing efficiency for detection of flaws by permeability systems are:

- Heat treated, quenched, drawn, and machine straightened
- Cold drawn and machine straightened
- Cold drawn, annealed, and machine straightened
- Hot rolled and machine straightened, centerless ground
- Hot rolled and machine straightened

After straightening, the bars should be aged 24 to 48 hours at near room temperature for optimum sensitivity of flaw detection. Aging can be hastened by stress relieving at low temperature in a furnace (up to 260 °C, or 500 °F).

The minimum seam depth that can be detected in cold drawn, straightened round bars is approximately 0.025 mm (0.001 in.) for each 1.6 mm ($1/16$ in.) of bar diameter; hexagonal and square bars with the same processing will be more sensitive. For example, in a 25 mm (1 in.) diameter round bar, a 0.41 mm (0.016 in.) seam is readily detected, while a 0.30 to 0.33 mm (0.012 to 0.013 in.) seam can be detected in hexagonal or square

bars. The reason for this difference lies in the residual stress levels imparted by the rotary and planar straighteners.

Other flaws, such as laps, slivers, cracks, hard or soft spots, dimensional changes, cupping, chevrons, and pipe, are readily indicated. For subsurface type flaws, detection is possible only if they lie within the normal penetration range and are of sufficient size to affect the inherent stress level. The penetration is approximately 6.4 mm (¼ in.) for low carbon steels, 7.9 mm (5/16 in.) for medium carbon steels, and up to 13 mm (½ in.) for many alloy steels.

One other factor not to be overlooked is the end effect, which prevents end-to-end inspection of the bar. As the front and rear ends of the bar enter and leave the magnetic field, the field is grossly distorted, preventing inspection of the end portions of the bar. For the average inspection speed of 37 to 46 m/min (120 to 150 sfm), the noninspected lengths will be as follows:

Bar diameter		Noninspected length at each end	
mm	**in.**	**mm**	**in.**
6.4–13	¼ – ½	102–152	4–6
13–25	½ – 1	152–203	6–8
25–50	1–2	203–305	8–12
50–75	2–3	305–406	12–16

The signal obtained for a flaw of given size, as well as the amount of end effect, will vary somewhat with the amount of draft used in drawing the bar. Using the normal 0.8 mm (1/32 in.) draft as the basis for comparison, a 1.6 mm (1/16 in.) draft will increase the signal size by 50%, while a 3.2 mm (⅛ in.) draft will produce an increase of about 90% (Fig. 19).

All the above values hold true only when the secondary test coil is of the proper size; that is, the inside diameter of the coil should be 3.2 to 6.4 mm (⅛ to ¼ in.) greater than the bar diameter. The diameters of bar

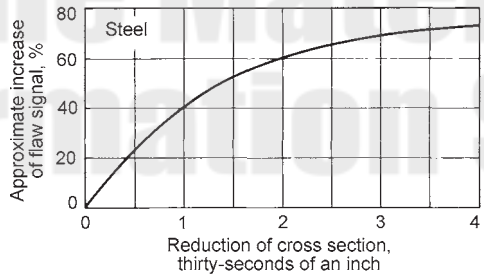


Fig. 19 Relationship between increase of flaw signal and increasing reduction of cross-section (increasing draft) for cold drawn steel bars. Base reference is a hot rolled bar. Source: Ref 1

stock inspected by these systems generally range from 4.8 to 140 mm ($^{3}/_{15}$ to 5½ in.).

Equipment for Detecting Flaws. The circuitry may include three types of electronic systems: the null system for the detection of flaws (as previously explained and shown in Fig. 18) and two identical standard systems, one of which is used for detecting mixed grades in a given lot of steel and the other for indicating variations of hardness or structure. All systems are independent and provide simultaneous indications with a single pass of the bar through the coil.

The null system utilizes a pair of matched windings that provides for the comparison of a section of the bar with another section spaced some distance from the first. The matched windings are connected in opposition, and the resultant voltage is theoretically zero, making the wave displayed on the oscilloscope a straight line. In practice, however, such a balance is seldom obtained. A small voltage with the wave shape showing two peaks phase displaced 180° can normally be seen on the oscilloscope screen (bar out, Fig. 20). The wave pattern changes when a bar is placed within coils (bar in, Fig. 20). Should a flaw of minimum depth be present, the change in the waveshape is too small for measurement, even though there is a differential voltage between the null coils. Therefore, other relationships must be used to provide the desired information.

The use of an electronic gate of any desired width permits these measurements to be made in any section of the wave. For example, the test gate shown in Fig. 20 is adjustable to any position of the 360° cycle. It is normally positioned 8 to 20° on either side of the stress peaks, where experience has revealed the wild stress effects are minimal and waveform changes for flaws are readily detectable. Most systems provide a second electronic gate that can monitor the section of wave shape where flaws cause a change in the saturation level, if this can be reached for the size and grade of material under test, deflections greater than a predetermined amount will energize a signal that indicates rejection.

Use for Sorting. The two standard systems differ from the null in that only one coil winding for each is utilized on the bar being tested (Fig. 18). The voltage derived from this coil is balanced by a voltage in the instrument that is fully adjustable to the degree that the zero-center meters can be adjusted to their midpoint while the oscilloscope presentation continually shows the distorted wave shape. Should any undesired bars appear within the lot being tested, the meter deflection will then provide power for activation of suitable alarm. The selectivity of the section of waveshape to be monitored is provided by an electronic gate, adjustable through 180°. Only half of the full 360° wave is required, the remainder being the negative duplicate of the positive and not shown on the oscilloscope. Both standard coil systems (Fig. 18) are fully independent and should be operated at different positions of the waveform to obtain as much information

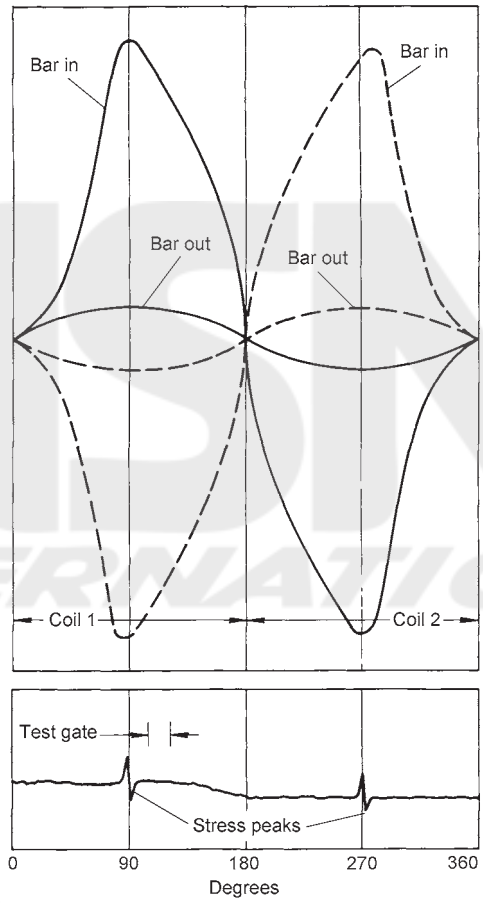


Fig. 20 Wave shape for oscilloscope pattern of a full electrical cycle for empty coils (bar out) and loaded coils (bar in). The position of an electronic gate for viewing an established portion of the cycle is shown. Source: Ref 1

as possible during the test. The standard system is used to monitor each bar in a lot for composition, hardness, structure, and size, and to indicate the presence of uniform depth seams, cracks, and laps, which generally escape detection by the null system.

In addition to coil arrangements such as those illustrated in Fig. 17 and 18, a fairly elaborate set of electronic gear is required for inspecting steel bars. Some type of equipment for handling the bars and conveying them through the coils at the desired rate is also required. The degree of sophistication designed into the equipment depends mainly on the number of similar bars to be inspected. Typical control units are adaptable to either the eddy current or the magnetic permeability systems of inspection. Many variations are commercially available.

ACKNOWLEDGMENT

This chapter was adapted from Nondestructive Inspection of Steel Bar, Wire, and Billets, *Nondestructive Evaluation and Quality Control*, Volume 17, *ASM Handbook*, 1989.

REFERENCES

1. Nondestructive Inspection of Steel Bar, Wire, and Billets, *Nondestructive Evaluation and Quality Control*, Vol 17, *ASM Handbook*, ASM International, 1989, p 549–560
2. N. Matsubara, H. Yamaguchi, T. Hiroshima, T. Sakamoto, and S. Matsumoto, Nondestructive Testing of Cold Drawn Wires and Cold Forged Products, *Wire J. Int.*, March 1986

CHAPTER 14

Inspection of Tubular Products

WROUGHT TUBULAR PRODUCTS are nondestructively inspected chiefly by eddy current techniques (including the magnetic flux leakage technique) and by ultrasonic techniques. In general, the eddy current and magnetic flux leakage techniques are applied to products not exceeding 1020 mm (40 in.) in diameter or 19 mm (¾ in.) in wall thickness. On the other hand, ultrasonic inspection is used on tubes ranging from 3.2 to 2030 mm (⅛ to 80 in.) in diameter and from 0.25 to 64 mm (0.01 to 2½ in.) in wall thickness. However, there are many exceptions, and the range of special techniques and applications associated with each inspection method is large. Most welded and seamless tubular products are nondestructively inspected by the manufacturer at the mill.

The many uses to which steel tubular products have been applied form a basis for classifying steel tubular products; for example, the terms used for the first classification *casing*, *tube*, and *pipe* are assigned on the basis of usage, as in water well casing, oil well tubing, and drill pipe. A second classification is based on methods of manufacture. Accordingly, all steel tubular products can be classified as either welded or seamless. A third classification applicable to special shapes can be considered subordinate to both of the general classifications above.

The major applications of the nondestructive inspection of tubular products are:

- Detection and evaluation of flaws
- Sorting of mixed stock
- Measurement of dimensions
- Comparative measurement of specific physical and mechanical properties

Of these, the primary application is the detection and evaluation of flaws. Sorting is often an auxiliary application employed for grade or size verification and can be based on chemical composition, dimensions, physical and mechanical properties, or other significant variables. A difficulty encountered in sorting arises when variables of little or no interest affect instrument indications to a greater degree than do the variables of interest.

Selection of Inspection Method

The fundamental factors that should be considered in selecting a nondestructive inspection method and in selecting from among the commercially available inspection equipment, are the product characteristics, nature of the flaws, extraneous variables, rate of inspection, end effect, mill versus laboratory inspection, specification requirements, equipment costs, and operating costs.

Product Characteristics. Among the product characteristics that may affect the choice of inspection method and equipment are tube or pipe diameter, wall thickness, surface condition, method of fabrication, electrical conductivity, metallurgical condition, magnetic properties (notably permeability), and degree of magnetization.

Nature of Flaws. Both the nature of flaws and of potential but unallowable deviations from certain specified dimensions or properties have a bearing on the selection of inspection methods and equipment. The nature of flaws is often markedly influenced by the method of manufacture. For example, flaws in welded pipe are usually confined to the vicinity of the weld; therefore, an inspection procedure that is confined to the weld area may be adequate. If the welds are resistance welds, the most usual flaws are located in the weld plane and are in effect two-dimensional, having length and width but negligible thickness. On the other hand, if the welds are arc welds, porosity is the most usual flaw. In all welded tubular products, cracks are the most damaging flaws. In seamless tube, the location of flaws is not restricted, but may occur anywhere in or on the tube section.

Extraneous Variables. Many of the measurable variables in tubing and pipe are normal to the product and are not cause for rejection. These extraneous or harmless factors sometimes exert a greater effect on the inspection instrument than do the flaws that must be detected. For example, variations in magnetic permeability are common in steel and generate large signals in instruments that are permeability sensitive. However, the signals often are not pertinent to the test, nor are they cause for rejection. Surface scratches may be cause for rejection in some products and yet may be acceptable in others. Consequently, the inspection method and instrument selected must ignore or minimize variables that will not affect the utility of the part in its intended application.

The rate of inspection required is a major factor in the selection of an inspection procedure. When the value of the part or the hazard associated

with its application justifies slow and thorough inspection, the procedure chosen is likely to be radically different from that selected for a mass produced, low cost product used in a noncritical application.

End Effect. In some applications, the only portions of the tube that are genuinely critical in its ultimate application are the ends. Unfortunately, with many nondestructive testing instruments, specific problems arise when inspection of the ends is required. End effect is encountered with the eddy current, ultrasonic, and radiographic methods. Consequently, inspection of the entire tube and the ends may require two different procedures; as a result, production speed is reduced and the total cost of inspection is correspondingly increased.

Mill Versus Laboratory Inspection. Although laboratory demonstrations of nondestructive inspection techniques may yield excellent results, subsequent mill performance may be entirely unsatisfactory because of conditions present in the mill that were not present in the laboratory.

Specification requirements may also affect the choice of inspection method and equipment. When the tubular product is covered by a flaw size specification all tubes with flaws larger than those specified must be rejected. However, tubes with flaws smaller than the specified rejection level should be accepted. Because many nondestructive inspection systems do not provide for linear adjustment or are incapable of making the required differentiation, this aspect of instrument performance must be carefully investigated.

Equipment cost is usually a major factor in the selection of inspection method and equipment. The initial cost of equipment may occasionally be minor, but in some cases installation may cost over $1 million in basic and related equipment. The lowest cost equipment may be for magnetic particle or liquid penetrant inspection. High cost installations involve automatic flaw marking, classification of product based on flaw magnitude, computer analysis of results, multiple sorting levels, and many other convenience factors.

The operating cost of inspection procedures and equipment varies widely. In general, it is inversely proportional to the cost of the installation. The more expensive installations are usually completely automatic and are incorporated in a production line whose primary function is something other than inspection. Consequently, inspection adds little to the total operating cost. In contrast, the lower cost installations usually involve a separate operation and require the services of a highly trained, skilled operator.

Inspection of Resistance Welded Steel Tubing

Resistance Welded Steel Tubing

The diameters of resistance (longitudinal) welded steel tubing range from about 13 to 914 mm (½ to 36 in.); wall thicknesses range from 0.38

to 13 mm (0.015 to 0.5 in.). Tubing of intermediate and smaller diameters is produced on a draw bench.

Flaws that occur in resistance welded steel tubing include cold welds, contact marks, cracks, pinholes, and stitching. The terminology used to designate such flaws varies; the terms used in this chapter are those adopted by the American Petroleum Institute (API).

Cold weld is the term widely used to indicate inadequate or brittle bonding with no apparent discontinuity in the fracture. Cold weld cannot be detected reliably by any nondestructive inspection method currently available.

Contact marks (electrode burns) (Fig. 1a) are intermittent imperfections near the weld line that result from miniature arcs between the welding electrode and the surface of the tube.

Hook cracks (upturned fiber flaws) (Fig. 1b) are separations within the base metal due to imperfections in the strip edge, which are parallel to

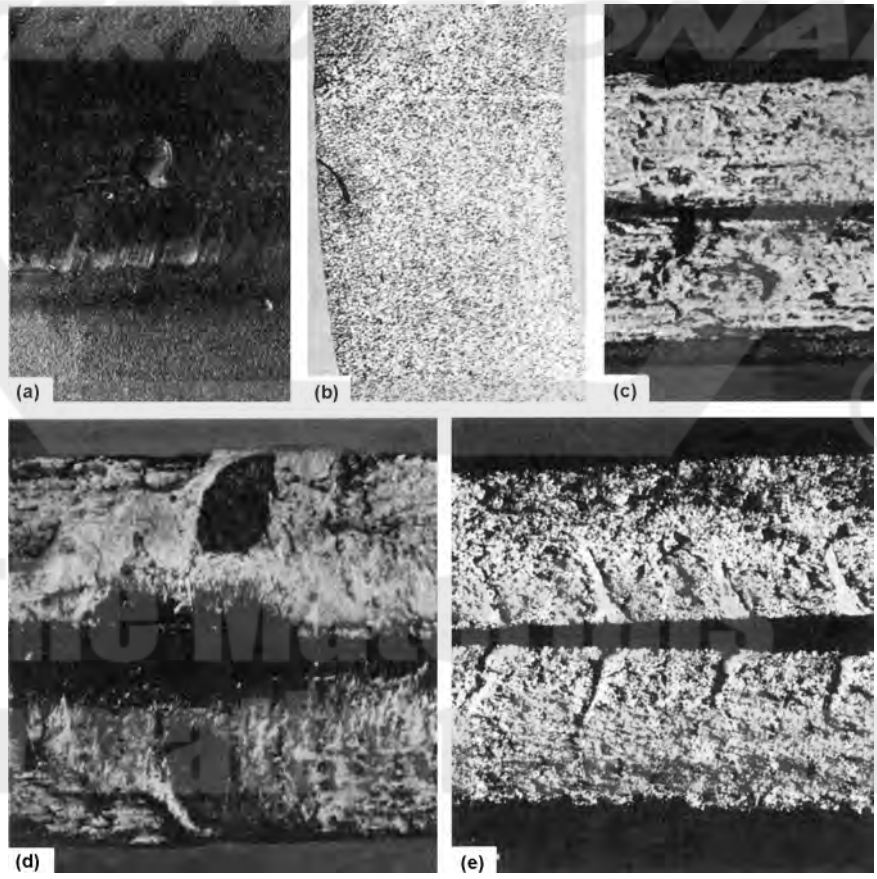


Fig. 1 Typical flaws in resistance welded steel tubing, (a) contact marks (electrode burns), (b) hook cracks (upturned fiber flaws), (c) weld area crack, (d) pinhole, (e) stitching. Views (c), (d), and (e) are mating fracture surfaces of welds. Source: Ref 1

the surface and turn toward the outside or inside surface when the edges are upset during welding.

Weld area cracks (Fig. 1c) are any cracks in the weld area not due to upturned fibers.

Pinholes (Fig. 1d) are minute holes located in the weld line.

Stitching (Fig. 1e) comprises a regular pattern of light and dark areas that are visible when the weld is broken in the weld line. The frequency of variation usually corresponds to weld-current variation. Increased use of ultrahigh frequency current for welding has minimized the occurrence of stitching.

The nondestructive inspection of resistance welded tube can be performed continuously on a welding machine or on individual lengths at any stage of processing. When performed on a welding machine, test indications can be used to guide the welding machine operator in making machine adjustments.

Eddy Current Inspection

Eddy current methods are probably the most widely used for the inspection of welded steel tubing in diameters up to 75 mm (3 in.), although these methods are not limited to the smaller diameters. Typical weld imperfections detectable by eddy current inspection are shown in Fig. 2. In welded tubing, most flaws occur in or near the longitudinal welded seam, and in most cases a test of a narrow band including the seam is adequate. This makes possible the use of small eddy current probe coils tangent to the seam area, eliminating the diameter limitation. Tubes with diameters up to 406 mm (16 in.) are currently being inspected by this method.

Eddy current inspection is usually performed on tubing having wall thicknesses less than 3.2 mm (⅛ in.), but successful production testing has been reported on tubing having wall thicknesses to 13 mm (½ in.). Most eddy current tests use differential systems and are most sensitive to flaws that involve a marked change in normal electrical characteristics. If the flaw is of considerable length and of uniform characteristics, it is sometimes necessary to use special arrangements for its detection. Small probe coils continuously compare the weld zone with the base metal, thus revealing the existence of the elongated or long flaw.

When inspecting for shallow crack-like surface flaws 3.2 mm (⅛ in.) or less in depth, relatively close correlation between crack depth and signal magnitude has been obtained with a single coil arrangement without magnetic saturation. However, the limited penetration of this arrangement and its need for a surface opening for depth evaluation limit its usefulness.

The speed of inspection by eddy current methods depends in part on many factors, including the size of the flaw that must be detected, the discriminating ability of the circuit used, end inspection requirements, and the speed of response of the signal circuit. The mathematical relationship

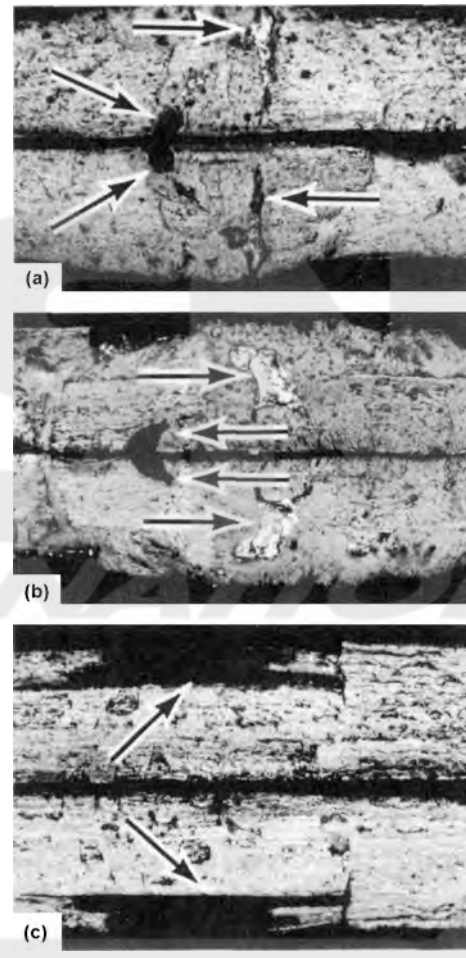


Fig. 2 Mating fracture surfaces of pipe or tube welds showing imperfections detectable by eddy current inspection, (a) unwelded spot (diagonal arrows) and a nonpenetrating pinhole (horizontal arrows); (b) unwelded spots, probably caused by entrapped foreign matter; (c) surface crack in weld. Source: Ref 1

of the test current frequency and linear speed may automatically limit the size of flaw that can be detected. Speeds of 305 m/min (1000 sfm) have been recorded, but the usual speed is 45 to 90 m/min (150 to 300 sfm).

Weld Twist. When using probe coils in eddy current inspection, the twist in the weld sometimes causes a special problem. When the weld twists out of the zone of high sensitivity, the effectiveness of flaw detection is markedly reduced. The problem cannot be solved by increasing the size of the arc segment covered by the probe coil, because this arrangement also reduces sensitivity. One solution involves the use of a series of small probe coils, staggered with respect to the weld line, to ensure continuous coverage. The problem has also been solved in some installations by taking advantage of the electromagnetic difference that exists between

the weld zone and the base metal. Special probe coils respond to this difference and automatically rotate the test head or the tube until the weld zone is properly located with respect to the probe coil.

End effect, caused by abrupt changes in the magnetic field, becomes a problem whenever cut lengths are inspected. Various auxiliary circuits, ranging widely in effectiveness, have been developed for suppressing end effect to permit satisfactory inspection closer to the end of the tube. End effect can be minimized by keeping the tube ends in contact as they move through the test coils.

Mechanical variables that may affect inspection results include transverse movement of the tube in the test coil and changes in temperature or linear speed. The contribution of these factors to test results is sometimes difficult to determine in the laboratory, but they may create serious problems in production testing.

Equipment costs for eddy current inspection can vary widely, depending on the extent of refined circuitry, automatic handling and sorting equipment, computer analyzers, or special auxiliary equipment that may be needed.

The operating costs of a well designed eddy current system are among the lowest of any nondestructive inspection method. After the system has been properly adjusted, it can be operated by unskilled workers. When automatic marking is provided, the inspection can frequently be combined with another operation without appreciably increasing the cost of the latter operation.

Advantages and Limitations. All flaws in resistance welds except cold weld are readily detected by eddy current methods. Cold weld is by far the most difficult of all flaws to detect by any of the nondestructive inspection methods.

Although the other types of flaws discussed can be detected by eddy current methods, it should not be inferred that all eddy current instruments will detect all of these flaws. The range of capabilities of commercial eddy current instruments is extensive, and conclusions regarding their capabilities often require actual tests. Because eddy current test coils may either surround or be adjacent to the tube being tested, the variety of coil designs, arrangements, and combinations constitutes another major group of variables affecting equipment capabilities. In general, eddy current instruments have the advantages of speed in testing and convenience in operating, marking, and sorting. Perhaps their most universal disadvantage is their inability to inspect completely to the ends of tubes.

Flux Leakage Inspection

Flux leakage (or magnetic field perturbation) inspection is similar to eddy current inspection but requires magnetization of the tube and is limited to the inspection of ferromagnetic materials. When the tube is magne-

tized to near saturation, the magnetic flux passing through the flaw zone is diverted by the flaws. Detectors of various types detect the diverted flux when either the detector or the tube is moved in a direction that causes the detector to cut through the diverted flux. This in turn produces a signal to reveal the presence of the flaw.

Various means are used to magnetize the tube. A current carrying conductor inside the tube produces a circular magnetic field, magnetizing the tube in a circumferential direction. The magnetic flux is diverted by the longitudinal component of any flaws in its path. The probe, moving through the diverted flux, generates a signal roughly proportional to the size of the flaw. On a longitudinal welded seam, an electromagnet with pole pieces on each side of the weld can be used to magnetize the weld area, with flux passing transversely across the seam. The magnetic flux is diverted by the longitudinal component of any flaw in the weld, and the flaw can be detected electronically. To detect transverse flaws, the tube may be magnetized longitudinally by an encircling conductor. The flux is then diverted by the transverse component of any flaw present, and the probe moving through the diverted or leakage flux reveals the presence of flaws.

Hall probes are the detectors ordinarily used. In all applications, there must be relative movement between the probes and the diverted flux so as to generate a signal and to indicate the presence of a flaw. The relative motion can be achieved by rotating or oscillating either the tube or the probes. As in eddy current inspection, various types of instrumentation have been developed and are available commercially.

Because of the nature of the flux leakage test, tube diameter is not a limitation, but the wall thickness that can be tested is limited by the ability of the magnetic flux to penetrate the wall and the ability of the sensor to sense flaws at a distance from the wall. Production applications have been used on tubing having wall thicknesses up to 25 mm (1 in.), but 7.6 mm (0.3 in.) is the usual limit. At wall thicknesses in excess of 7.6 mm (0.3 in.), sensitivity becomes a serious problem.

Although the flux leakage method usually detects flaws that are longitudinally oriented, the principle of the flux leakage method can be used in the design of equipment for detecting transverse flaws. Pinholes, with minimal longitudinal dimensions, and subsurface flaws are difficult to detect. For reliable detection of isolated pinholes, the pitch of the helical inspection path must be small, and the production rate is correspondingly limited. Sensitivity to subsurface flaws drops rapidly as the flaws are located farther from the surface. To detect inside surface flaws such as cracks and gouges, flux leakage equipment requires special design features for reliable quantitative evaluation.

The speed of inspection is a function of the dimensions of the elements involved and the maximum tolerable length of flaw. Because the tube or the probe must be rotated or oscillated, only a helical or zigzag path is in-

spected, and the pitch of the helix or the distance between reversals must be less than the maximum tolerable length of flaw. When the tube must be fed over a central conductor for magnetization and then removed for inspection, high speed production is hindered. The use of multiple probes reduces the actual testing time in proportion to the number of probes, but the time required feeding the tube over the conductor remains constant. Such installations are operated in production at speeds as high as 15 m/min (50 sfm). Installations using external magnetization are reported to operate at speeds to 60 m/min (200 sfm).

Weld twist presents a problem in any installation in which only the weld is inspected. In flux leakage inspection, the problem is solved by increasing the magnitude of the arc covered by the oscillating probe. As in eddy current inspection, the error caused by end effect can be minimized by butting the ends of the tubes together during the test. Mechanical conditions, such as tube ovality, variations in linear speed, and transverse movement of the tube, have adverse effects on the test results and must be controlled.

As with eddy current equipment, the cost of equipment for flux leakage inspection varies. An elementary unit costs slightly more than a comparable eddy current unit because of the need for rotating devices. The addition of auxiliary equipment, such as automatic markers, recorders, computer analyzers, and special handling devices, markedly increases cost.

Operating costs, which are relatively low, depend on the degree of automation and the degree to which the inspection can be combined with other operations. Flux leakage tests can sometimes be combined with another operation, for example, welding. As the tube emerges from the welding operation, it enters the field of the electromagnetic yoke (Fig. 3), which generates a flux in the weld area. The oscillating probe detects any flux diverted by a flaw in the weld. However, in most cases, the need for move-

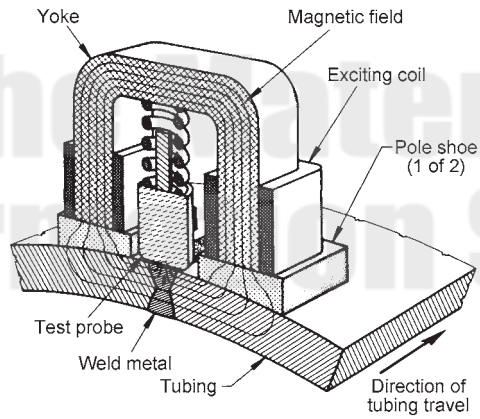


Fig. 3 Setup for the flux leakage inspection of welded steel tubing. Source: Ref 1

ment of the probe through the diverted flux makes the combination less desirable than systems with no moving parts.

Ultrasonic Inspection

Ultrasonic inspection is one of the most widely used methods for inspecting tubular products. Widespread use of ultrasonics on tubular products was made practical by the development of angle beam shear wave testing, immersion testing, and focused transducers. As with the eddy current and flux leakage methods, ultrasonic inspection can be applied either to the entire tube or to the weld only.

Ultrasonic inspection of the entire welded tube is usually limited to small diameter, drawn products, in which the weld cannot easily be distinguished from the remainder of the tube. The tubing may have a diameter as small as 3.2 mm (⅛ in.) and a wall thickness of only 0.25 mm (0.01 in.). These small products are usually inspected while immersed in water (immersion inspection). They are rotated as they pass longitudinally through a glanded immersion tank. The immersed transducers must be carefully selected for tube diameter, wall thickness, and type of imperfection to be located. The transducer, focal length, response to outside diameter and inside diameter calibration notches, instrumentation pulse rate, gate adjustment for flaw alarm, and speed of tube travel are all variables to be taken into consideration. Inspection is usually performed slowly (0.9 m/min or 3 sfm). Tubes must be clean, straight, round, and of uniform dimensions. All types of flaws that commonly occur in resistance welds, except cold weld, can be detected by ultrasonic inspection.

Most ultrasonic inspection of resistance welded tubing is restricted to the weld zone and is performed immediately after the welding operation. Components and adjustments for inspecting the weld must be carefully selected and accurately controlled. The transducer must be appropriate for the size and type of flaws to be detected. Focused transducers are generally preferred. The shear wave angle must be selected for the best evaluation of imperfections. The angle often used is 45°, but tests have revealed that angles between 50 and 70° yield signals more directly proportional to the area of flaws in the weld plane.

In the inspection of pipe, provision must be made to maintain the spacing between the transducer and the pipe constant within close tolerances as the pipe moves past the transducer. The couplant should preferably be continuously delivered to the surface of the pipe through openings in the transducer mounting. Coupling through a water jet is also used. Particular attention should be given to the detection of short flaws. Some ultrasonic pipe inspection equipment will not detect flaws shorter than 6.4 mm (¼ in.), which will not satisfy the inspection requirements for most resistance welded pipe.

A disadvantage of the ultrasonic method in tube inspection is its high sensitivity to minor scratches and to elongated dimensional changes, such as the ridge left when the weld flash is not completely removed or rolled down. However, proper selection of inspection equipment can minimize this problem. An important development is the wheel type search unit. The transducer of the wheel type search unit is mounted on the axle of a liquid filled wheel and is held in a fixed position as the wheel rotates. The surface of the wheel is flexible and adapts itself to the surface condition of the tube as it rolls over it. A small amount of liquid couplant, usually water, is required between the surface of the wheel tire and the surface of the tube. This arrangement provides most of the advantages of immersion testing without the necessity of immersing the tube.

The speed of inspection is limited by the pulse rate of the ultrasonic equipment and by the maximum length of a tolerable imperfection. Speeds as high as 69 m/min (225 sfm) have been reported, but unless multiple inspection heads are used speed is ultimately dependent on the rejectable flaw size.

Weld twist can present a problem; as the weld twists away from the critical location, transducer sensitivity drops sharply. To maintain the weld and the transducer in the correct mechanical relationship, the weld can be positioned automatically by the use of an electromagnetic control.

End effect, although less of a problem than in eddy current inspection, is a factor in ultrasonic inspection; and supplementary testing may be necessary if inspection of the tube ends is critical. The supplementary test can be made with ultrasonic equipment of special design. Mechanical variables are critical in contact ultrasonic testing. Spacing between the transducer and the surface of the tube, angle of transducer, and sidewise movement of tube must be accurately controlled. These variables can sometimes be better controlled in immersion testing.

The equipment costs of ultrasonic inspection equipment are highly dependent on the amount of auxiliary equipment included. Accessories such as automatic marking devices, computer analyzers, and material handling equipment can markedly increase equipment costs, especially for the inspection of heavy pipe.

The operating costs of ultrasonic inspection, in accord with other inspection methods, depend on whether inspection is operated separately or combined with another operation. For example, if inspection is incorporated into the welding line, an inspector usually is not required, and the operating costs are minimal.

Magnetic Particle Inspection

The principal use of the magnetic particle method in the inspection of resistance welded pipe is largely limited to the inspection of pipe ends. In

some pipe applications, the ends of the pipe are the sections most critically loaded, and magnetic particle inspection of the ends supplements inspection of the remainder of the pipe by other methods. In the past, the method was widely used to inspect the entire area. However, its inability to detect significant subsurface flaws, even when the magnetic particles are coated with a fluorescent, and its dependence on human vision and judgment led to its replacement by eddy current and ultrasonic methods. The magnetic particle method is still used in the mill to help establish the precise location of flaws previously detected by other inspection methods.

Liquid Penetrant Inspection

Liquid penetrants (visible dye and fluorescent) are ordinarily used on nonferromagnetic materials, which constitute only a small fraction of resistance welded tubular products. Testing speeds are extremely slow, and use of these methods can be justified only when the hazard involved in end use justifies extreme inspection precautions. In such cases, the penetrant methods usually supplement other methods.

Radiographic Inspection

Radiographic methods of inspection cannot be used successfully on the longitudinal seam of resistance welded pipe, because the predominant flaws are essentially two-dimensional and have little or no effect on the radiographic film. However, when the ends of resistance welded pipe are butt welded together, arc welding is frequently used, and the method normally used to inspect arc welded joints is radiography.

Seamless Steel Tubular Products

Steels melted by various processes can be successfully converted into seamless tubes. In general, killed steels made by open hearth, electric furnace, and basic oxygen processes are used. Because of the severity of the forging operation involved in piercing, the steels used for seamless tubes must have good characteristics with respect to both surface and internal soundness. A sound, dense cross-section, free from center porosity or ingot pattern, is the most satisfactory for seamless tubes. Metallurgical developments have contributed greatly to the improvement of steels for seamless tubes. As a result, the seamless process has been extended to include practically all of the carbon and alloy grades of steel.

Flaws in seamless tubular products may occur at any point on the outside and inside surfaces or within the tube wall. The flaws usually encountered as illustrated in Fig. 4 are:

- *Blisters* (Fig. 4a) are raised spots on the surface of the pipe caused by the expansion of gas in a cavity within the wall.

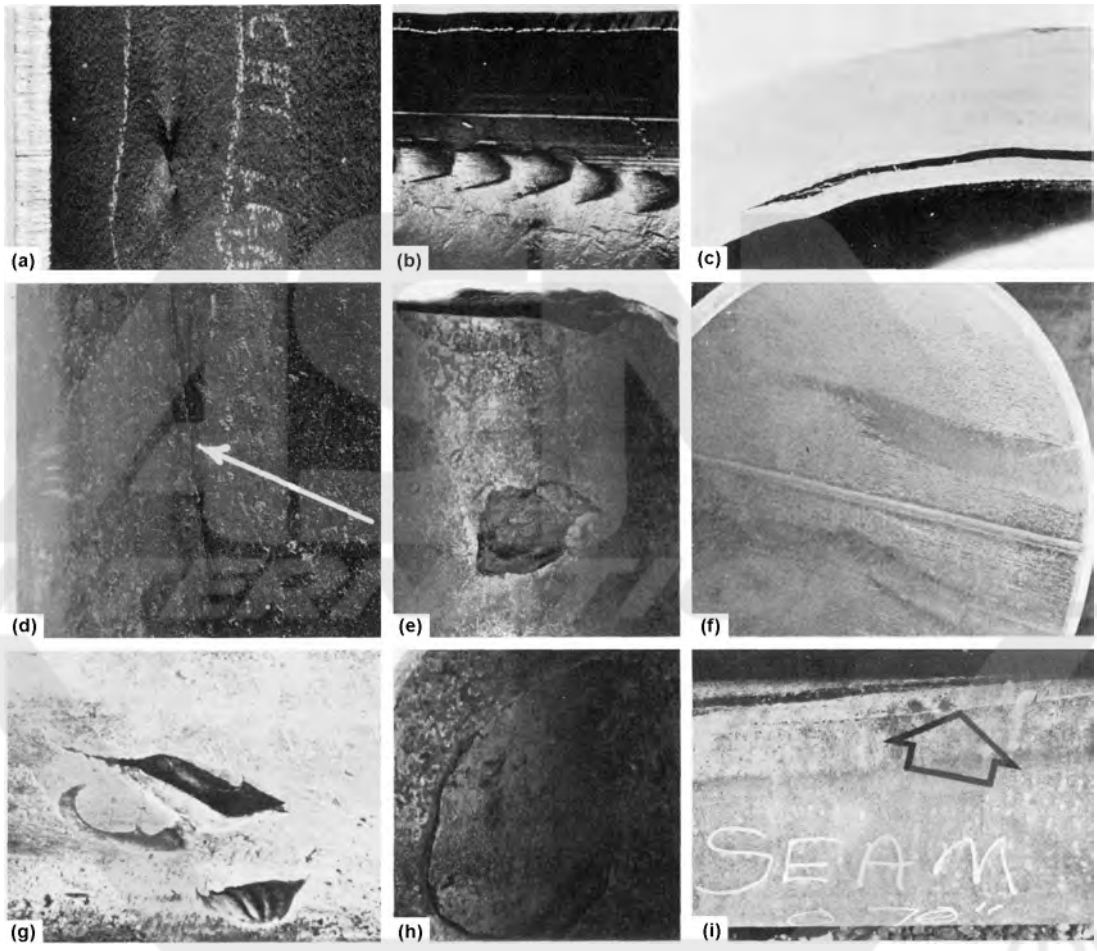


Fig. 4 Typical flaws in seamless tubing, (a) blister, (b) gouge, (c) lamination, (d) lap (arrow), (e) pit, (f) plug scores, (g) rolled-in slugs, (h) scab, (i) seam (arrow). Source: Ref 1

- *Gouges* (Fig. 4b) are elongated grooves or cavities caused by the mechanical removal of metal.
- *Laminations* (Fig. 4c) are internal metal separations creating layers generally parallel to the surface.
- *Laps* (Fig. 4d) are folds of metal that have been rolled or otherwise worked against the surface but that have not been fused into sound metal.
- *Pits* (Fig. 4e) are depressions resulting from the removal of foreign material rolled into the surface during manufacture.
- *Plug scores* (Fig. 4f) are internal longitudinal grooves, usually caused by hard pieces of metal adhering to the mandrel, or plug, during plug rolling.
- *Rolled-in slugs* (Fig. 4g) are foreign metallic bodies rolled into the metal surface, usually not fused.

- *Scabs* (Fig. 4h) are flaws in the form of a shell or veneer, generally attached to the surface by sound metal. Usually, scabs originate as ingot flaws.
- *Seams* (Fig. 4i) are crevices in rolled metal that have been closed by rolling or other work but have not been fused into sound metal.

Ultrasonic Inspection

Ultrasonic inspection is probably the method most commonly used on seamless tubular products, which range from 3.2 to 660 mm (1/8 to 26 in.) in diameter and from 0.25 to 64 mm (0.01 to 2 in.) in wall thickness. The tube, while rotating, is usually moved longitudinally past the transducers, thus providing inspection along a helical path. In a typical installation, six transducers inspect the rotating tube as it progresses through the machine. Four transducers are below the tube and are coupled to it by water columns; two transducers are above the tube and make contact through fluid filled plastic wheels. This machine is capable of handling tubes ranging from 50 to 305 mm (2 to 12 in.) in diameter.

As with welded tubing, the smaller diameters of seamless products are inspected while immersed, and the larger diameters are inspected by direct contact. Transducer crystals may be quartz, lithium sulfate, barium titanate, or lead zirconate. The frequency ordinarily used is 2.25 MHz. The transducer may or may not be focused, depending on the tube diameter and the nature of the flaws anticipated. In most cases, the ultrasonic shear wave angle is 45° but may be as large as 70°. The usual couplant is water, but oil is sometimes used. Because all installations involve rotation of either the tube or the transducers, the inspection invariably follows along a helical path. The pitch and width of the helical path vary widely, depending on the characteristics of the equipment and the specifications to be met. The pitch is usually between 9.5 and 13 mm (0.374 and ½ in.), which translates to two to three revolutions per each 25 mm (1 in.) of longitudinal travel. Almost all forms of visible and audible alarms, as well as automatic recorders, are used with the ultrasonic equipment.

All types of flaws in seamless tubes can be detected by ultrasonic methods, but the minimum flaw dimensions, the degree of sensitivity, the flexibility of adjustment, and the accuracy of calibration all vary widely with the basic instrumentation and the supplementary components chosen. The flaw used most frequently for calibration is a longitudinal slot. The depth of slot may vary from 3 to 12 % of the wall thickness, depending on the end use of the product and the specification involved. The length of the slot may be as much as 38 mm (1 in.) but is usually specified as twice the width of the transducer. Widths of slots should be kept at a minimum and should never exceed twice the depth. The frequency of calibration checks depends on the criticality of the tube application. In a few cases, a calibration check is required after every tube, but once every four hours of production is usually considered adequate.

Speed of inspection also varies widely, depending on many variables, especially the maximum tolerable length of flaw and the number of transducers used. The current range of inspection speed is 0.6 to 46 m/min (2 to 150 sfm); the upper limit can be increased by increasing the number of transducers.

Multiple search units can be used in the immersion tank to provide several simultaneous inspections during one pass of the pipe or tube through the unit. Specifications may call for circumferential inspection from two directions because the reflection from a flaw may vary, depending on the direction in which the ultrasonic beam strikes it. In addition to the circumferential inspection to detect longitudinal flaws, a longitudinal (axial) inspection may be required to detect transverse flaws. Also, it may be desirable to ultrasonically measure wall thickness, eccentricity, or both. All these tests can be performed at the same time by utilizing search units that are designed for the tests and that are properly positioned in the tank. Normally, rejection is based on the presence of flaw indications exceeding those from the reference notch. Reworking and reinspection are generally permitted if other requirements, such as minimum wall thickness, are satisfied. Other refinements are included in, or can be added to, the inspection system. For example, the feeding of tubes to the unit and their withdrawal can be automated. Various audible and visible alarm systems and marking devices can be added.

Normally, the water couplant used in the system is filtered and deaerated. Air entrapped in the water can produce false indications. Similarly, water on the inside surfaces of tubes will produce false signals, and these surfaces must be kept dry. Tubes are connected to each other by stoppers or by taping the ends together. The glands at each end of the water tank must be cut in a manner that will allow passage of the tubes without undue loss of water couplant. Finally, air must be prevented from being drawn into the entry gland along with the tube. This is accomplished by directing a stream of water over the outside of the tube just before it enters the gland.

Eddy Current Inspection

In eddy current inspection, use of an encircling detector coil is limited to a maximum tube diameter of about 75 mm (3 in.). As tube diameter increases, the ratio of flaw size to tube diameter decreases; consequently, the flaw is increasingly more difficult to detect. This problem is overcome by using several small probe coils and with spinning probes. When probe coils are used, the flaw becomes a significantly high percentage of the zone surveyed.

Because independently mounted probes ride over the tube surface, good magnetic coupling is ensured. Magnetic saturation is used to obtain maximum sensitivity to flaws close to, or on, the inside surface of the tube, and

the frequency of the test current is kept relatively low, sometimes as low as 1 kHz. Internal spinning probes can also be used if a lower production rate can be tolerated.

In some installations, the eddy current test with magnetic saturation is supplemented by a probe type eddy current test, which in effect provides high sensitivity inspection of the surface. Each of the four probe coils shown in Fig. 5, serves as both inductor and detector, and rotates about the tube as it moves longitudinally through the rotating assembly. Magnetic saturation is not required. The segment of the tube where the flaw occurs is identified by markers. In Fig. 6, the test head is shown ready for use, with eight paint spray guns in position for marking the proper zone.

Eddy current inspection is used on seamless tubular products ranging from 3.2 to 244 mm (⅛ to 9⅝ in.) in diameter and from 0.25 to 14.0 mm (0.01 to 0.55 in.) in wall thickness. In most cases, magnetic saturation is used, but when the primary concern is the detection of surface imperfections, small probe coils without magnetic saturation are used. If the steel is entirely nonmagnetic, no saturation is required in any system. The frequency used ranges from 1 to 400 kHz and depends on such variables as the wall thickness of the pipe or tube, the coil design and arrangement, and the use of saturation. The spacing between the pipe or tube and the test coil(s) varies widely. However, for high sensitivity and accuracy, it should be kept to a minimum and is occasionally kept to as little as 0.25

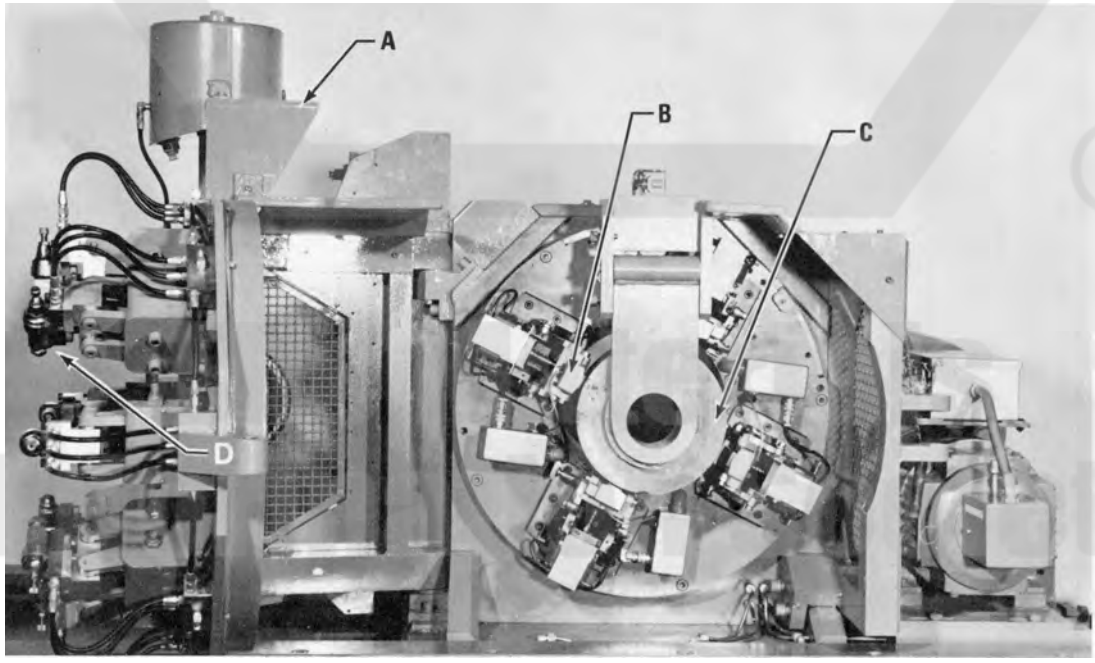


Fig. 5 Unit used for the probe type eddy current inspection of seamless steel tubing; A, outer cover, containing test head, in open position; B, one of four rotating eddy current probe coils; C, reference standard test piece in position for calibration; D, one of eight paint spray guns for marking. Source: Ref 1

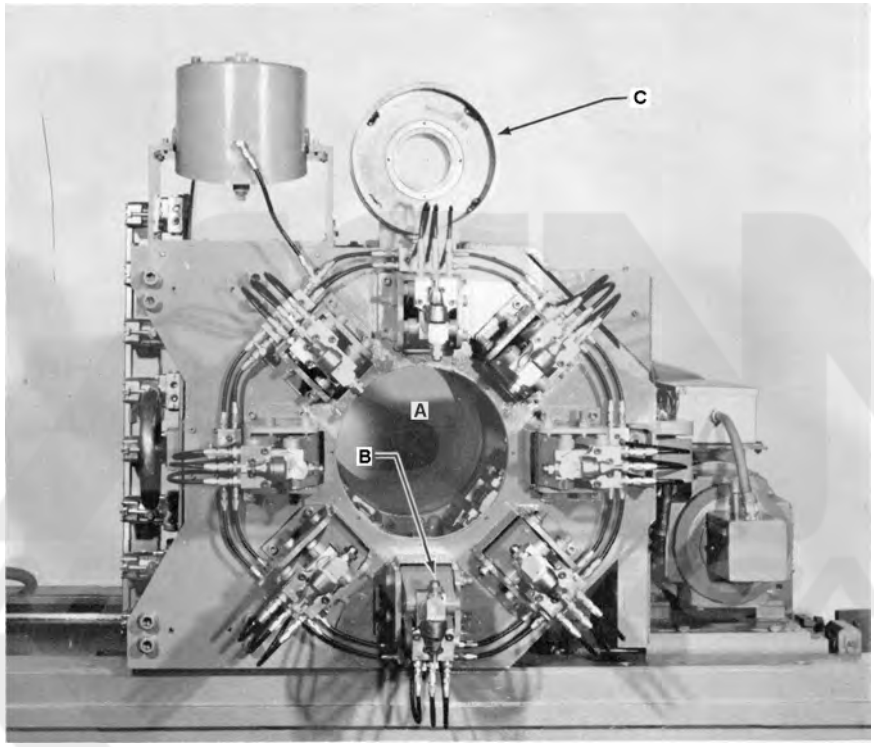


Fig. 6 Test head of the eddy current inspection unit shown in Fig. 5; A, orifice for test pipe or tube; B, one of eight paint spray guns for marking; C, reference standard test piece. Source: Ref 1

mm (0.01 in.). This clearance is insufficient for practical use, and the usual clearance for production testing ranges from 1.6 to 19 mm ($^1/_{16}$ to ¾ in.).

Although all flaws that usually occur in seamless pipe or tube can be detected by eddy current methods, external flaws are more easily detected than internal flaws. Laminations are the most difficult flaws to detect. Some installations are intended to detect surface flaws only.

Flux leakage techniques are used for the inspection of seamless tubular products ranging from 32 to 914 mm (1¼ to 36 in.) in diameter and from 3.2 to 19 mm (⅛ to ¾ in.) in wall thickness. Because the flaws usually have a significant longitudinal dimension, transverse magnetic fields are usually used. Longitudinal fields can be used but are rarely considered necessary. As a rule, the transverse magnetic field is produced by a current passing through a conductor located in the center of the pipe or tube. In some cases, the field is produced between the poles of an electromagnet or a permanent magnet whose pole pieces are shaped to fit the pipe or tube diameters. The signal generating movement of either the tube or the probe with respect to the other is accomplished by rotating either tube or probe.

Sensitivity to inside surface flaws is a problem when using these methods; the problem becomes more serious as the wall thickness increases. In

some cases, the solution to the problem is a rotating internal probe moving through the tube or kept stationary while the tube moves. Other installations depend on electronic filters and the difference in frequency between the signals for internal and surface flaws.

The production rate of flux leakage inspection depends on many factors. The maximum permissible speed of probe or tube movement with respect to the other is about 90 m/min (300 sfm). The circuits will respond almost instantaneously when inspection speeds are kept below this speed limit. The principal limiting factor in production output speed then becomes the maximum tolerable length of the flaw, which in turn governs the pitch of the helix inspected. However, the production rate can be increased to any desired level by the simultaneous inspection of several segments of the same pipe or tube. Actual inspection speeds range from 0.9 to 60 m/min (3 to 200 sfm), depending on the diameter, the system used, sensitivity required, and other variables. The methods that use external magnets can inspect at a much higher overall rate than those that depend on an internal conductor for magnetization.

Magnetic particle and liquid penetrant inspection methods are simple and economical and are most useful for surface inspection in specialized, small scale applications. When applied to welded tubing, their use can be restricted to the weld zone. However, when applied to seamless tubing, there are no surface restrictions. The inability of these methods to locate small flaws beneath the surface and their dependence on the vision, alertness, and judgment of the inspector limit their usefulness in meeting modern specifications for seamless steel tubing.

Radiographic Inspection. The principal application of radiography to seamless tubing, as with welded tubing, is the inspection of girth welds joining the ends of tubes. Even in this application, it should be supplemented by magnetic particle inspection.

Nonferrous Tubing

A wide variety of nonferrous alloy tubing, such as tubing made of brass, copper, aluminum, nickel, titanium, and zirconium, can be inspected for cracks, seams, splits, and other flaws. Eddy current inspection is the method most widely used, followed by the ultrasonic and liquid penetrant methods.

Eddy Current Inspection

When eddy current inspection is employed for nonferrous tubing, the range of tube diameters normally permits the use of an encircling coil. The typical flaws respond well to differential type coils. The frequencies employed usually range from 1 to 25 kHz. The choice of frequency is generally dependent on the electrical conductivity and wall thickness of the tub-

ing. Because magnetic saturation is not required, the inspection equipment is simpler and more compact than that used on ferromagnetic tubing. Tubes range from 3.2 to 89 mm (⅛ to 3½ in.) in diameter, with wall thicknesses from 0.25 to 14.2 mm (0.01 to 0.56 in.). Testing speeds to 370 m/min (1200 sfm) have been reported. On a limited basis, eddy current inspection is also being applied to finned copper tubing.

Immersion ultrasonic inspection is used on tubes ranging in diameter from 6.4 to 254 mm (¼ to 10 in.), with wall thicknesses as small as 0.25 mm (0.01 in.). In some installations, four channels are used, two for the detection of transverse flaws and two for longitudinal flaws. Because these tests are usually critical, they are performed at low speeds, usually less than 3 m/min (10 sfm).

The liquid penetrant inspection of nonferrous tubular products is performed in the conventional manner.

ACKNOWLEDGMENT

This chapter was adapted from Nondestructive Inspection of Tubular Products, *Nondestructive Evaluation and Quality Control*, Volume 17, *ASM Handbook*, 1992.

REFERENCES

1. Nondestructive Inspection of Tubular Products, *Nondestructive Evaluation and Quality Control*, Vol 17, *ASM Handbook*, ASM International, 1992, p 561–581

CHAPTER 15

Inspection of Forgings

IN FORGINGS of both ferrous and nonferrous metals, the flaws that most often occur are caused by conditions that exist in the ingot, by subsequent hot working of the ingot or the billet, and by hot or cold working during forging. The inspection methods most commonly used to detect these flaws include visual, magnetic particle, liquid penetrant, ultrasonic, eddy current, and radiographic inspection.

Flaws Originating in the Ingot

Many large open-die forgings are forged directly from ingots. Most closed-die forgings and upset forgings are produced from billets, rolled bar stock, or preforms. Many, though by no means all, of the imperfections found in forgings can be attributed to conditions that existed in the ingot, sometimes even when the ingot has undergone primary reduction prior to the forging operation. Some, but again by no means all, of the service problems that occur with forgings can be traced to imperfections originating in the ingot.

Chemical Segregation. The elements in a cast alloy are seldom distributed uniformly. Even unalloyed metals contain random amounts of various types of impurities in the form of tramp elements or dissolved gases that are seldom distributed uniformly. Therefore, the composition of the metal or alloy will vary from location to location.

Deviation from the mean composition at a particular location in a forging is termed *segregation*. In general, segregation is the result of solute rejection at the solidification interface during casting. For example, the gradation of composition with respect to the individual alloying elements exists from the cores of dendrites to interdendritic regions. Therefore, segregation produces a material having a range of compositions that do not have identical properties. Forging can partially correct the results of seg-

regation by recrystallizing or breaking up the grain structure to promote a more homogeneous substructure. However, the effects of a badly segregated ingot cannot be totally eliminated by forging; rather, the segregated regions tend to be altered by the working operation, as shown in Fig. 1.

In metals, the presence of localized regions that deviate from the nominal composition can affect corrosion resistance, forging and joining (welding) characteristics, mechanical properties, fracture toughness, and fatigue resistance. In heat treatable alloys, variations in composition can produce unexpected responses to heat treatments, which result in hard or soft spots, quench cracks, or other flaws. The degree of degradation depends on the alloy and on process variables. Most metallurgical processes are based on the assumption that the metal being processed is of a nominal and reasonably uniform composition.

Ingot Pipe and Centerline Shrinkage. A common imperfection in ingots is the shrinkage cavity, commonly known as pipe, often found in the upper portion of the ingot. Shrinkage occurs during freezing of the metal, and eventually there is insufficient liquid metal near the top end to feed the ingot. As a result, a cavity forms, usually approximating the shape of a cylinder or cone, hence the term pipe. Piping is illustrated in Fig. 2. In addition to the primary pipe near the top of the ingot, secondary regions of piping and centerline shrinkage may extend deeper into an ingot.

Primary piping is generally an economic concern because it is cropped before rolling or forging, but if it extends sufficiently deep into the ingot body and goes undetected, it can eventually result in a defective forging. Detection of the pipe can be obscured in some cases if bridging has occurred.

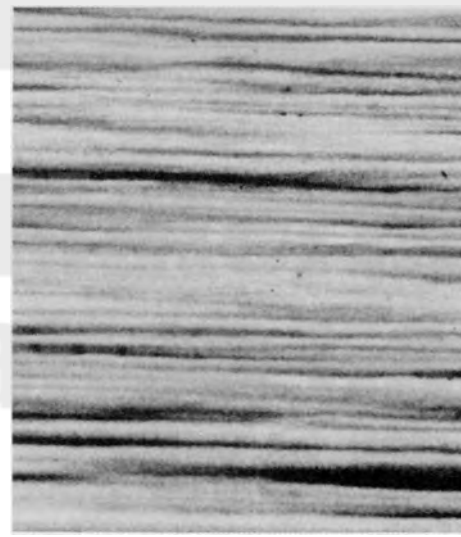


Fig. 1 Microstructural bonding due to chemical segregation and mechanical working. Source: Ref 1

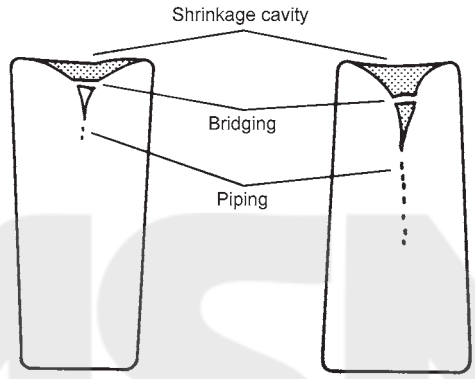


Fig. 2 Schematic showing piping in top poured ingots. Source: Ref 1

Piping can be minimized by pouring ingots with the big end up, by providing risers in the ingot top, and by applying sufficient hot top material (insulating refractories or exothermic materials) immediately after pouring. These techniques extend the time that the metal in the top regions of the ingot remains liquid, thus minimizing the shrinkage cavity produced in the top portion of the ingot.

On the other hand, secondary piping and centerline shrinkage can be very detrimental, because it is harder to detect at the mill and may subsequently produce centerline defects in bar and wrought products. Such a material condition may indeed provide the flaw or stress concentrator for a forging burst in some later processing operation or for a future product failure.

High Hydrogen Content. A major source of hydrogen in certain metals and alloys is the reaction of water vapor with the liquid metal at high temperatures. Water vapor may originate from the charge materials, slag ingredients and alloy additions, refractory linings, ingot molds, or even the atmosphere itself if steps are not taken to prevent such contamination. The resulting hydrogen goes into solution at elevated temperatures; but as the metal solidifies after pouring, the solubility of hydrogen decreases, and it becomes entrapped in the metal lattice.

Hydrogen concentration in excess of about 5 ppm has been associated with flaking, especially in heavy sections and high carbon steels. Hydrogen flakes (Fig. 3) are small cracks produced by hydrogen that has diffused to grain boundaries and other preferred sites, for example, at inclusion/matrix interfaces. However, hydrogen concentrations in excess of only 1 ppm have been related to the degradation of mechanical properties in high strength steels, especially ductility, impact behavior, and fracture toughness.

Metals can also possess high hydrogen content without the presence of flakes or voids. In this case, the hydrogen may cause embrittlement of the material along selective paths, which can drastically reduce the resistance

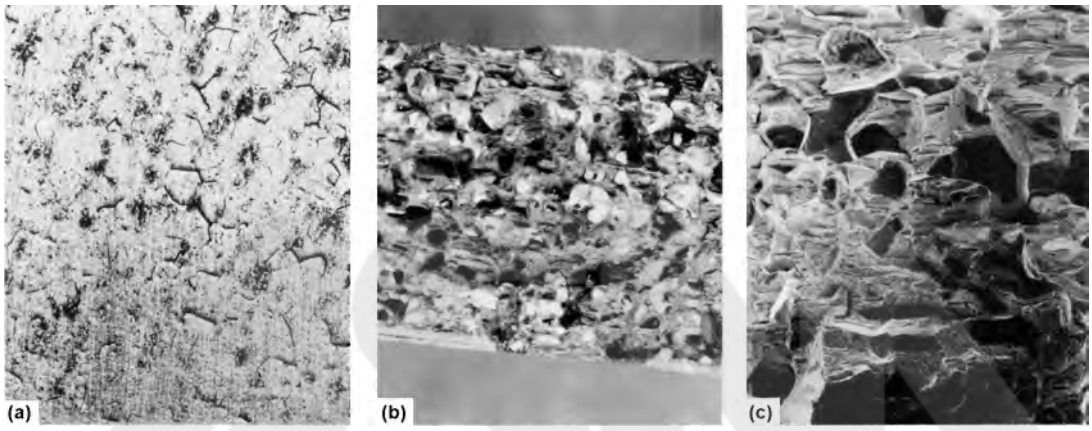


Fig. 3 Hydrogen flaking in an alloy steel bar. (a) Polished cross-section showing cracks due to flaking. (b) Fracture surface containing hydrogen flakes. Note the reflective, faceted nature of the fracture. (c) SEM micrograph showing the intergranular appearance of the flakes in this material. Source: Ref 1

of a forged part to crack propagation resulting from impact loading, fatigue, or stress corrosion.

In cases where hydrogen related defects can serve as the initiation site for cracking and thus increase the likelihood of future failures, it is advisable to use a thermal treatment that can alleviate this condition. For example, slow cooling immediately following a hot working operation or a separate annealing cycle will relieve residual stresses in addition to allowing hydrogen to diffuse to a more uniform distribution throughout the lattice and, more important, to diffuse out of the material.

Nonmetallic inclusions, which originate in the ingot, are likely to be carried over to the forgings, even though several intermediate hot working operations may be involved. Also, additional inclusions may develop in the billet or in subsequent forging stages.

Most nonmetallic inclusions originate during solidification from the initial melting operation. If no further consumable remelting cycles follow, as in air melted or vacuum induction products (with no remelting cycle to follow), the size, frequency, and distribution of the nonmetallic inclusions will not be altered or reduced in size or frequency during further processing. If a subsequent vacuum remelting operation is used, the inclusions will be lessened in size and frequency and will become more random in nature. If an electroslag remelting cycle is used, a more random distribution of inclusions will result.

The two kinds of nonmetallic inclusions that generally occur in metals are:

- Those that are entrapped in the metal inadvertently and originate almost exclusively from particles of matter that are occluded in the metal while it is molten or being cast

- Those that separate from the metal because of a change in temperature or composition

Inclusions of the latter type are produced by separation from the metal when it is in either the liquid or the solid state.

Oxides, sulfides, nitrides, or other nonmetallic compounds form droplets or particles when these compounds are produced in such amounts that their solubility in the matrix is exceeded. Air melted alloys commonly contain inclusions, mainly of these chemical characteristics. Vacuum or electroslag remelted alloys more commonly contain conglomerates of any of these types, frequently combined with carbon or the hardening element or elements that precipitate during stabilization and aging cycles to form inclusions such as titanium carbonitrides or carbides.

Homogenizing cycles are normally used for the ingot prior to conversion or at an early stage of conversion. Because these compounds are products of reactions within the metal, they are normal constituents of the metal, and conventional melting practices cannot completely eliminate them. However, it is desirable to keep the type and amount of inclusions to a minimum so that the metal is relatively free from those inclusions that cause the most problems.

Of the numerous types of flaws found in forgings, nonmetallic inclusions appear to contribute significantly to service failures, particularly in high integrity forgings such as those used in aerospace applications. In many applications, the presence of these inclusions decreases the ability of a metal to withstand high static loads, impact forces, cyclical or fatigue loading, and sometimes corrosion and stress corrosion. Nonmetallic inclusions can easily become stress concentrators because of their discontinuous nature and incompatibility with the surrounding composition. This combination may very well yield flaws of critical size that, under appropriate loading conditions, result in complete fracture of the forged part.

Unmelted electrodes and shelf are two other types of ingot flaws that can impair forgeability. Unmelted electrodes (Fig. 4a) are caused by chunks of electrodes being eroded away during consumable melting and dropping down into the molten material as a solid. Shelf (Fig. 4b) is a condition resulting from uneven solidification or cooling rates at the ingot surfaces.

The consumable melting operation has occasionally been continued to a point where a portion of the stinger rod is melted into the ingot, which may be undesirable because the composition of the stinger rod may differ from that of the alloy being melted. To prevent this occurrence, one practice is to weld a wire to the stinger rod, bend the wire down in tension, and weld the other end of the wire to the surface of the electrode a few inches below the junction of the stinger rod and the electrode. When the electrode has been consumed to where the wire is attached to it, the wire is released and springs out against the side of the crucible, thus serving as an alarm to

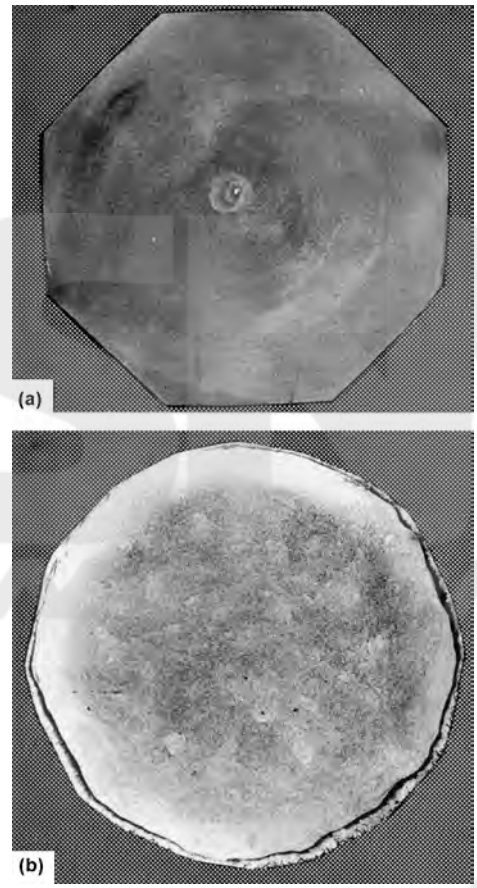


Fig. 4 Sections through two heat-resistant alloy ingots showing flaws that can impair forgeability. (a) Piece of unmelted consumable electrode (white spot near center). (b) Shelf (black line along edge) resulting from uneven solidification of the ingot. Source: Ref 1

stop the melting. A disadvantage of this practice is that the wire may become detached and contaminate the melt.

Flaws Caused by the Forging Operation

Flaws produced during forging (assuming a flaw-free billet or bar) are the result of improper setup or control. Proper control of heating for forging is necessary to prevent excessive scale, decarburization, overheating, or burning. Excessive scale, in addition to causing metal loss, can result in forgings with pitted surfaces. The pitted surfaces are caused by the scale being hammered into the surface and may result in unacceptable forgings.

Severe overheating causes burning that results in the melting of the lower melting point constituents. This melting action severely reduces the

mechanical properties of the metal, and the damage is irreparable. Detection and sorting of forgings that have been burned during heating can be extremely difficult.

In many cases, the flaws that occur during forging are the same as, or at least similar to, those that may occur during hot working of the ingots or billets; these are described in the previous section.

Internal flaws often appear as cracks or tears, and they can result either from forging with too light a hammer or from continuing forging after the metal has cooled down below a safe forging temperature. Bursts can also occur during the forging operation.

A number of surface flaws can be produced by the forging operation. These flaws are often caused by the movement of metal over or upon another surface without actual welding or fusing of the surfaces; such flaws may be laps or folds.

Cold shuts often occur in closed-die forgings. They are junctures of two adjoining surfaces caused by incomplete metal fill and incomplete fusion of the surfaces.

Surface flaws weaken forgings and can usually be eliminated by correct die design; proper heating; and, correct sequencing and positioning of the workpieces in the dies.

Shear cracks often occur in steel forgings; they are diagonal cracks occurring on the trimmed edges and are caused by shear stresses. Proper design and condition of trimming dies to remove forging flash are required for the prevention of shear cracks.

Other flaws in steel forgings that can be produced by improper die design or maintenance are internal cracks and splits. If the material is moved abnormally during forging, these flaws may be formed without any evidence on the surface of the forging.

Selection of Inspection Method

The principal factors that influence the selection of a nondestructive inspection (NDI) method for forgings include degree of required integrity of the forging, metal composition, size and shape of the forging, and cost. There are sometimes other influential factors, such as the type of forging method used.

For high integrity forgings, it is often required that more than one inspection method be employed because some inspection methods are capable of locating only surface flaws; therefore, one or more additional methods are required for locating internal flaws. For example, many forgings for aerospace applications are inspected with liquid penetrants (or with magnetic particles, depending on the metal composition) for locating surface flaws, then by ultrasonics for detecting internal flaws.

Certain conditions unique to forgings can create service problems, yet they are not easily detected by nondestructive inspection. Exposed end

grain, which can lead to poor corrosion resistance or to susceptibility to stress corrosion cracking, is the most prevalent of the undesirable conditions. The strength of the forging can be adversely affected if the grains flow in an undesirable direction (as in grain reversal) or if grain flow is confined to only a portion of the section being forged, rather than being well distributed. Both exposed end grain and poor grain flow can be most effectively corrected by redesigning the forging or forging blank, particularly with regard to flash. It is virtually impossible to detect exposed end grain nondestructively; only rarely can poor grain flow be detected nondestructively. Both conditions are much more readily analyzed by sectioning and macroetching sample forgings in the preproduction stage of product development.

When certain steels or nonferrous alloys are forged at too high a temperature, or sometimes when a part cools too slowly after forging, there is the potential for excessive grain growth. Such a condition is difficult to detect nondestructively, except with ultrasonics, and then only when the grains are very large. Even with very large grains, ultrasonic inspection cannot determine grain size quantitatively, nor can it detect large grains reliably. The possibility that large grains are present can only be inferred from excessive attenuation of the ultrasonic beam.

Effect of Type of Forging

Many of the types of flaws that can occur in forgings do so without particular regard to the type of forging; that is, open die, closed-die, upset, or rolled. However, there are many cases in which a specific type of flaw is more likely to occur in one type of forging than in another.

Open-Die Forgings. Most forgings produced in open dies are relatively large; therefore, their size is likely to impose some restrictions not only on the inspection method used but also on the system within a given inspection method. For large open-die forgings, NDI methods (other than visual) are generally limited to magnetic particle or liquid penetrant inspection (for surface discontinuities) or to ultrasonic inspection (for internal flaws). In general, the flaws likely to be found in open-die forgings are similar to those that may occur in other hot-worked shapes, with the exception of forging laps and cold shuts, which usually occur only in closed-die forgings.

Closed-Die and Upset Forgings. The discontinuities in closed-die forgings that can be detected by liquid penetrant inspection or magnetic particle inspection (if the forging is ferromagnetic) are:

- Forging laps, which can be caused by incorrect die design, use of incorrect size of forging stock, excessive local conditioning of forging stock for removal of surface flaws, and excessively sharp corners in the forging stock

- Seams, due to incomplete removal of seams from the forging stock
- Surface cracks, caused by incorrect forging temperature, nonductile metallic or nonmetallic segregates in the forging stock, or surface contamination from the furnace atmosphere or other contaminants in the furnace or on the forging (such as high sulfur fuels for heating nickel alloys or leaded crayons used for marking parts before heating)
- Quench cracks
- Cold straightening cracks

The likelihood that any of the listed discontinuities will appear in a closed-die forging produced in a press or by upsetting is more prevalent than for hammer forgings, because press and upset forgings permit no opportunity to monitor the workpiece being forged during the forging operation. During hammer forging, the top surface of the forging is visible between hammer blows. The bottom surface may also be visible at times–particularly for large forgings that are raised intermittently during forging for descaling and lubricating the bottom surface or for forgings of temperature sensitive alloys that are raised off the bottom die to permit heat recovery to the bottom surface.

Any seam in the forging stock or incipient laps or cracks will probably develop into significant forging laps or cracks if not detected during formation. Consequently, in hammer forging, a large percentage of such potentially scrap forgings can be removed from the production run and can either be salvaged by removing the discontinuity prior to finishing or scrapped at that point to avoid wasted forging time. Also, multiple cavity hammer forgings permit inspection of the parts and blending out of minor laps or superficial cracks before finish forging.

Most discontinuities detected by ultrasonic inspection and not detectable by either magnetic particle or liquid penetrant inspection are:

- Flakes, due to the absorption of hydrogen
- Pipe, due to center shrinkage in the ingot and subsequent insufficient reduction of forging stock
- Subsurface nonmetallic segregation
- Subsurface cracking, which may occur in certain alloys, particularly during the forging of irregular sections
- Weak centers in forging stock, caused by insufficient reduction from the ingot
- Subsurface cracks caused by forging material having comparatively cold centers, and generally occurring in large forging billets heated for insufficient time
- Rewelded forging laps, formed and rewelded during forging. With subsequent hammer blows, the lap forms, the scale is knocked or blown off, and the lapped metal rewelds, forming a healed lap with transverse grain flow and possibly entrapped scale

The presence of a rewelded forging lap in a suspected area can be checked by removing the surface metal below the decarburized layer and polishing this surface and swabbing with cold ammonium persulfate, thus revealing the decarburization at both sides of the lap (if present). The condition can be eliminated with corrections in blocker die design.

Ring-Rolled Forgings. Discontinuities in forgings produced by ring-rolling can either be inherited from the ingot or mechanically induced during forging, much the same as for forgings produced by hammers or presses.

Inherited discontinuities are common to all products produced from bar or billet and can usually be traced back to the composition, cleanliness, or condition of the ingot. Although these discontinuities are not found only in ring-rolled forgings, they probably account for the vast majority of known discontinuities in ring-rolled forgings. Typical discontinuities are inclusions, porosity, hot-top remnants, and segregation.

Ultrasonic inspection is a reliable method of detecting the presence of inherited discontinuities. It is always advisable to inspect the material before it is ring-rolled. Extremely large billets (120 cm or 48 in. in diameter and larger) may have surface conditions that cause problems relative to sound entry. Large billets may exhibit structural conditions, depending on the amount of reduction from the ingot, that are too large for a complete ultrasonic inspection. Smaller, well worked billets can be examined at 2.25 MHz. The larger billets require the use of 1.0 MHz crystals, and even then ultrasonic penetration is not always possible.

Final proof that the forging is free of inherited discontinuities is accomplished through ultrasonic inspection of the completed forging. Depending on the final machined shape, certain ring-rolled forgings may require a preliminary ultrasonic inspection before final machining. When an extreme change in contour prevents a complete ultrasonic inspection of the final shape, inspection can be performed on portions of the forged ring.

Externally induced mechanical discontinuities that have been found in ring-rolled forgings include surface related laps, cracks, and exfoliations. Normally, these discontinuities can be detected visually either during the manufacturing process or in the machined condition after rolling. However, ultrasonic inspection can be valuable for determining the depth of surface related discontinuities and for detecting them even when they have been obscured by subsequent working or metal movement. Magnetic particle inspection is also used to detect surface related, externally induced mechanical discontinuities in ferromagnetic ring-rolled forgings. Liquid penetrant inspection has often been successfully used to detect surface flaws in nonferromagnetic rings.

Mechanically induced internal discontinuities also known as strain induced porosity can occur in certain materials. Some nickel or titanium alloys have an inherent susceptibility to these types of discontinuities in portions of the ring stretched at critical temperatures. This can result from

improper stock distribution, improper rolling techniques, or improper tooling. In extreme cases, in which the induced porosity is excessive, the rupturing can progress to one or more external surfaces. Either form of internal mechanically induced discontinuity can initiate in an area where inherited discontinuities are present.

Effect of Forging Material

Some types of forging flaws are unique to specific work metals, and may influence the choice of inspection method.

Steel Forgings. The most common surface flaws in steel forgings are seams, laps, and slivers. Other surface flaws include rolled-in scale, ferrite fingers, fins, overfills, and underfills. The most common internal flaws found in steel forgings are pipe, segregation, nonmetallic inclusions, and stringers.

Either magnetic particle or liquid penetrant inspection can be used for steel forgings, although magnetic particle inspection is usually preferred. Only liquid penetrant inspection can be used for some stainless steel or nonferrous forgings. The selection of an inspection method depends on the size and shape of the forging and on whether the forging can be moved to the inspection station or the inspection equipment can be moved to the forging. For either inspection method, systems are available for inspecting forgings of almost unlimited size and weight. In most cases, magnetic particle inspection is less expensive and faster than liquid penetrant inspection.

Heat-Resistant Alloy Forgings. Most of the flaws found in heat-resistant forgings are those related to scrap selection, melting, or primary conversion to bar or billet, or those that occur during forging or heat treatment. Tramp elements, such as lead or zinc, have been present in scrap charges at levels that have caused hot shortness and a degradation of hot tensile ductility occurring near 370 °C (700 °F) of air melted alloys. No NDI method can reliably evaluate the presence or absence of possible tramp element contamination, and composition checks or hot tensile checks are too random for complete assurance of the presence or absence of contamination. The positive corrective action is to use a vacuum remelted product.

Melt related discontinuities, such as inclusions, pipe, unhealed center conditions, flakes, or voids, are the types of discontinuities that most frequently exist in heat-resistant alloy forgings. Ultrasonic inspection can detect and isolate these conditions. Segregated structures, unmelted electrodes, or portions of stringer rods are types of discontinuities that may sometimes be found in heat-resistant alloys.

Flakes (internal cracks) can be produced each time the material is heated and cooled to room temperature. The random orientation does not always present a properly oriented reflector for ultrasonic inspection, but

in most cases flaking can be detected ultrasonically with a high degree of reliability.

Unmelted pieces of electrode or shelf conditions appear infrequently in vacuum melted alloys. Either of these conditions can seriously degrade forgeability. Macroetching or another appropriate type of surface inspection of the machined forging or the billet is the most effective method of detecting unmelted electrodes or shelf conditions.

Seams are common and are readily detected by visual, magnetic particle, or liquid penetrant inspection. Grinding cracks are caused by severe grinding, which promotes network-type cracking on the surface of the material being conditioned. The network-type cracking may be present immediately after grinding or may not occur until subsequent heating for further forging. Seams and grinding cracks will cause severe surface rupturing during forging.

Center bursts occur during conversion to bar or billet if reduction rates are too severe or temperatures are incorrect; they are readily detected by ultrasonic inspection.

Ingot pipe, unhealed center conditions, or voids are melt-related discontinuities, but their occurrence in forgings is often a function of reduction ratio. The conversion practice must impart sufficient homogenization or healing to produce a product with sound center conditions. An example of an unsound condition that did not heal is shown in Fig. 5. Macroetching and ultrasonic inspection methods are the most widely used for identifying regions of unsoundness.

The nickel-base heat-resistant alloys are highly susceptible to surface contamination during heating for forging. Fuel oils containing sulfur will induce a grain boundary attack, which will cause subsequent rupturing during forging. Paint or marking crayons with high levels of similar contaminants will cause similar areas of grain boundary contamination.

Surface contamination is not normally detected by NDI methods prior to heating and processing, but if present at a level high enough to cause contamination, rupturing during forging will occur that can be detected by visual inspection. If the contamination occurs after final forging, with no subsequent metal deformation, the contaminated areas will be apparent as areas of intergranular attack. Macroetching, followed by liquid penetrant inspection, should be used.

Advanced forging processes, such as isothermal and hot-die forging, and the increasing use of computer modeling have greatly reduced problems associated with heat-resistant alloy forgings.

Nickel Alloy Forgings. The discontinuities that occur in nickel alloy forgings are generally of the same type as those found in heat-resistant alloy forgings; namely, cracks (external and internal), tears, seams, laps, coarse grain wrinkles, inclusions, and pipe. Although all metals may be subject to thermal cracking during forging, the age hardenable nickel al-

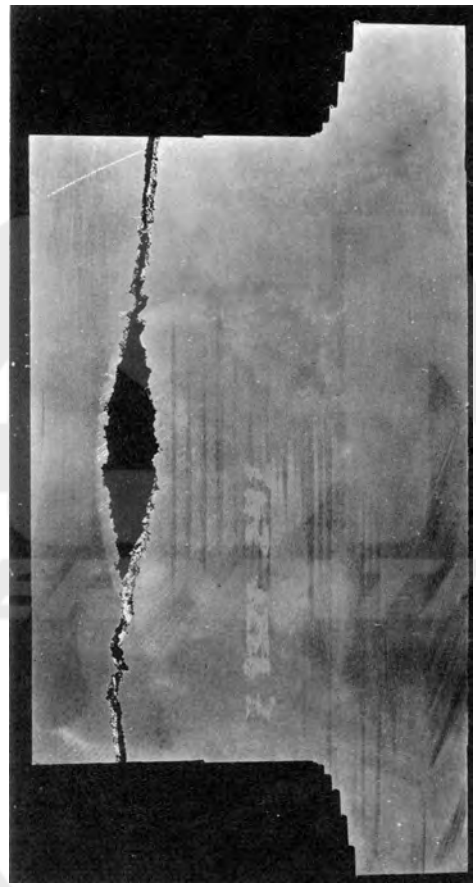


Fig. 5 Section through a heat-resistant alloy forging showing a central discontinuity that resulted from insufficient homogenization during conversion. Step machining was used to reveal the location of the rupture; original diameter is at right. Source: Ref 1

loys are more vulnerable than most other metals, thus requiring close temperature control during forging to avoid large temperature gradients.

Internal discontinuities in nickel alloy forgings can be located by ultrasonic inspection. Liquid penetrants are most often used to inspect for surface flaws; magnetic particles can be used if the alloy is sufficiently magnetic.

Aluminum Alloy Forgings. Common surface discontinuities in aluminum alloy forgings are laps, folds, chops, cracks, flow-throughs, and suckins (Fig. 6 and 7). The generation of these discontinuities is associated with the forging operation, processing practices, or design. Cracks can also result from seams in the forging stock.

The internal discontinuities that occur in aluminum alloy forgings are ruptures, cracks, inclusions, segregation, and occasionally porosity. Rup-

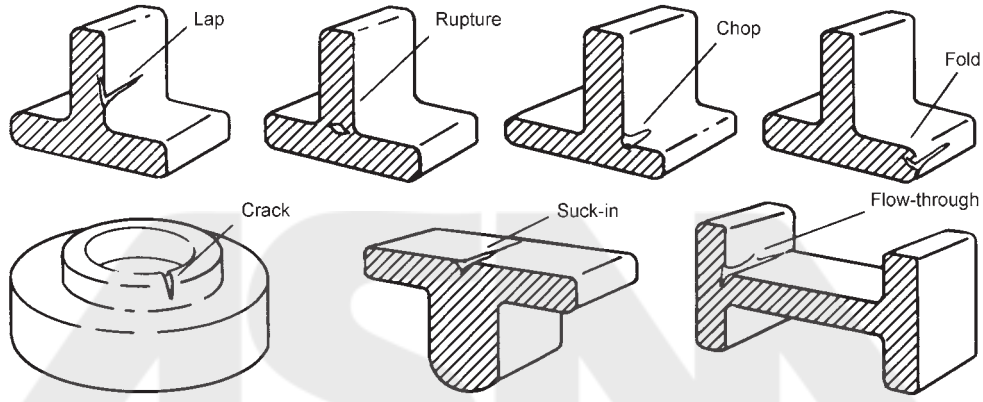


Fig. 6 Typical discontinuities found in aluminum alloy forgings. Source: Ref 1

Fig. 7 Band of shrinkage cavities and internal cracks in a 7075-T6 forging. The cracks developed from the cavities, which were produced during solidification of the ingot and which remained during forging because of inadequate cropping. Etched with Keller's reagent. Original magnification 9×. Source: Ref 1

tures and cracks are associated with temperature control during preheating or forging or with excessive reduction during a single forging operation. Cracks can also occur in stock that has been excessively reduced in one operation. Inclusions, segregation, and porosity result from forging stock that contains these types of discontinuities.

The inspection of aluminum alloy forgings takes two forms: in-process inspection and final inspection. In-process inspection, using techniques such as statistical process control and/or statistical quality control, is used to determine if the product being manufactured meets critical characteristics and if the forging process is under control. Final inspection, including

mechanical property testing, is used to verify if the completed forging product conforms to all drawing and specification criteria. Typical final inspection procedures used for aluminum alloy forgings include dimensional checks, heat treatment verification, and nondestructive evaluation.

Dimensional Inspection. All final forgings are subjected to dimensional verification. For open-die forgings, final dimensional inspection may include verification of all required dimensions on each forging or the use of statistical sampling plans for groups or lots of forgings. For closed-die forgings, conformance of the die cavities to the drawing requirements, a critical element in dimensional control, is accomplished prior to placing the dies in service by using layout inspection of plaster or plastic casts of the cavities. With the availability of computer-aided design (CAD) databases on forgings, such layout inspections can be accomplished more expediently with computer-aided manufacturing (CAM) driven equipment, such as coordinate measuring machines or other automated inspection techniques. With verification of die cavity dimensions prior to use, final part dimensional inspection may be limited to verifying the critical dimension controlled by the process (such as die closure) and monitoring the changes in the die cavity. Further, with high definition and precision aluminum forgings, CAD databases and automated inspection equipment, such as coordinate measuring machines and two-dimensional fiber optics, can be used in many cases for actual part dimensional verification.

Heat Treatment Verification. Proper heat treatment of aluminum alloy forgings is verified by hardness measurements and, in the case of 7*xxx*-T7*xxx* alloys, by eddy current inspection. In addition to these inspections, mechanical property tests are conducted on forgings to verify conformance to specifications. Mechanical property tests vary from the destruction of forgings to tests of extensions and/or prolongations forged integrally with the parts.

Nondestructive Inspection. Aluminum alloy forgings are frequently subjected to nondestructive inspection to verify surface or internal quality. The surface finish of aluminum forgings after forging and caustic cleaning is generally good.

A root mean square (rms) surface finish of 3.2 μm (125 μin.) or better is considered normal for forged and etched aluminum alloys; under closely controlled production conditions, surfaces smoother than 3.2 μm (125 μin.) rms can be obtained. Selection of NDI requirements depends on the final application of the forging. Surface quality is verified by liquid penetrant, eddy current, and other techniques when required. Aluminum alloy forgings used in aerospace applications are frequently inspected for internal quality using ultrasonic inspection techniques.

Magnesium alloy forgings are subject to the same types of surface and internal discontinuities as aluminum alloy forgings. In addition, surface cracks are common in magnesium alloy forgings and are usually caused by insufficient control of the forging temperature.

Visual inspection and liquid penetrant inspection are used to detect surface discontinuities. Ultrasonic inspection is used to locate internal discontinuities.

Titanium Alloy Forgings. Discontinuities that are most likely to occur in titanium alloy forgings are usually carried over into the bar or billet. Typical discontinuities in titanium alloy forgings are α-stabilized voids, macrostructural defects, unsealed center conditions, clean voids, and forging imperfections.

Alpha-stabilized voids are among the most common discontinuities found in forgings of titanium alloys. Research has determined that voids surrounded by oxygen-stabilized grains may be present in the ingot (Fig. 8). Because of the size of these voids and the coarse grain nature of the ingot, they cannot be detected until the ingot has been suitably reduced in cross-section and refined in structure. When the structure has been refined, the voids can be detected by ultrasonic inspection. Also, when the section is reduced sufficiently, radiographic inspection can be effectively used.

Alpha voids do not readily deform during forging, nor do they align with the flow pattern, as do typical inclusions in carbon or alloy steels. In most cases, α voids appear to be somewhat globular. Extremely small voids do not present an especially ideal target or reflector for ultrasonic energy. Attempts to correlate size with amplitude of indication obtained during ultrasonic inspections have not been completely reliable. For critical application forgings, the material is most often inspected twice–once in the bar or billet form before forging and again after forging. Because forging further refines structure and reorients possible discontinuities in

Fig. 8 Ti-8Al-Mo-1V, as-forged. Ingot void (black), surrounded by a layer of oxygen stabilized α (light). The remaining structure consists of elongated α grains in a dark matrix of transformed β. Etched with Kroll's reagent (ASTM 192). Original magnification 25×. Source: Ref 1

relation to the sound entry surface, the forging operation probably enhances the probability of detecting these discontinuities.

Macrodefects. Three principal defects are commonly found in macrosections of ingot, forged billet, or other semifinished product forms. These include high-aluminum defects (Type II defects), high-interstitial defects (Type I defects or low-density interstitial defects), and β flecks. High-aluminum defects are areas containing an abnormally high amount of aluminum. These are soft areas in the material (Fig. 9) and are also referred to as α segregation. Defects referred to as β segregation are sometimes associated with segregation. These are areas in which aluminum is depleted. The high interstitial defects (Fig. 10) are normally high in oxygen and/or nitrogen, which stabilize the α phase. These defects are hard and brittle; they are normally associated with porosity, as illustrated in Fig. 8.

Fig. 9 Ti-6Al-4V α-β processed billet illustrating the macroscopic appearance of a high aluminum defect. Original magnification 1.25×. Source: Ref 1

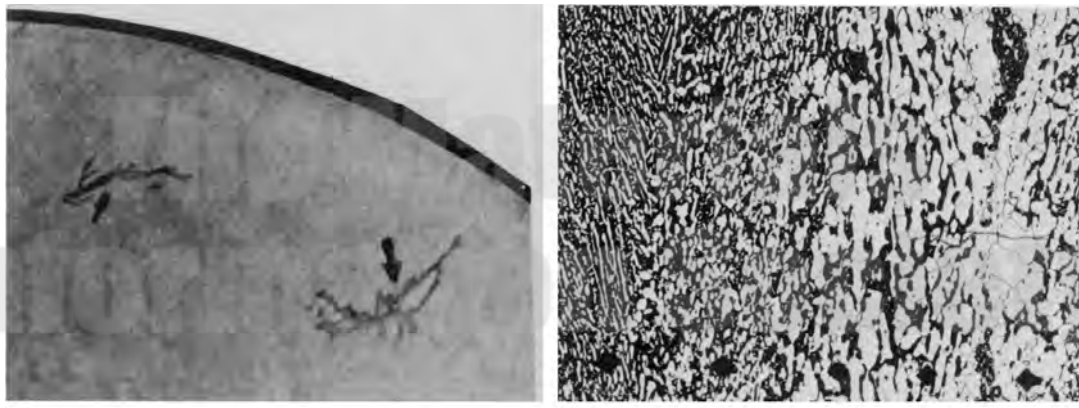


Fig. 10 Macrodefects in titanium billets. (a) Ti-6Al-4V α-β processed billet illustrating macroscopic appearance of a high interstitial defect, actual size. (b) Original maginification 100×. The high oxygen content results in a region of coarser and more brittle oxygen α stabilized than observed in the bulk material. Source: Ref 1

Beta flecks are regions enriched in a β-stabilizing element due to segregation during ingot solidification. Figure 11 shows the macroscopic appearance of β flecks in a Ti-6Al-6V-2Sn forging billet.

Unsealed center conditions are associated with insufficient ingot reduction. These are more prevalent in the larger stock sizes (>230 mm or 9 in. in diameter) and are normally removed by adequate croppage at the mill. Clean voids describe a condition that can be associated with unsatisfactory sealing of porosity elsewhere in the ingot or through center porosity formed during ingot reduction.

Nondestructive Inspection. Ultrasonic inspection is the best method of inspecting titanium alloy forgings. Inspection techniques are normally tailored to the rejection level indicated in the specifications and to the physical condition of the material being inspected. Surface conditions usually must be ideal; i.e., grain size must be fine and structural conditions must be controlled. Most airframe or similar static parts are inspected with equipment settings based on a No. 3 flat bottom hole standard. For the examination of critical rotating forgings for aircraft gas turbine engines, it is not uncommon to inspect to the equivalent of a No. 1 flat bottom hole standard. Experience with these highly critical forgings that rotate at high speeds in the presence of extreme temperature and pressure has indicated that small voids can initiate cracks and have caused catastrophic failures.

For satisfactory ultrasonic inspection of forgings to these stringent requirements, special techniques and equipment are usually required. Specially designed ultrasonic electronic equipment is used with focused or otherwise unique transducers. Also required are an immersion tank with rotating devices, automatic small incremental indexing devices, and automatic alarms for signal level. Special reference blocks are required, along with the usual flat bottom hole reference blocks. The correct indexing increment must be established, the linear alignment of the ultrasonic unit must be verified, and calibration checks must be made. All information must be recorded and retained for future reference.

Fig. 11 Ti-6Al-6V-2Sn α-β forged billet illustrating macroscopic appearance of β flecks that appear as dark spots. Etched with 8 mL HF, 10 mL HF, 82 mL H_2O, then 18 g/L (2.4 oz/gal.) of NH_4HF_2 in H_2O. Less than 1×. Source: Ref 1

Visual Inspection

Despite the many sophisticated inspection methods available, unaided visual inspection is still important and is often the sole method of inspecting forgings used for common hardware items. Under proper lighting conditions, the trained eye can detect several types of surface imperfections, including certain laps, folds, and seams. Visual inspection is often used first then questionable forgings are further examined by macroetching and inspection with macrophotography or some type of nondestructive method.

The only equipment necessary for visual inspection is a bench on which to place the forging and suitable cranes or hoists for forgings that are too heavy to lift by hand. Good and well controlled lighting conditions are essential. Optical aids such as magnifying glasses that can magnify up to about ten diameters are often used to increase the effectiveness of visual inspection.

Magnetic Particle Inspection

Magnetic particle inspection is useful for detecting surface imperfections as well as certain subsurface imperfections that are within approximately 3 mm (⅛ in.) of the surfaces in forgings of steel, some grades of stainless steel, and other ferromagnetic metals. Magnetic particle inspection can be used with fluorescent particles and ultraviolet light.

The advantages of magnetic particle inspection are:

- Almost instant results can be obtained in locating surface and certain subsurface imperfections
- Equipment can be transported to the forging, or the forging can be transported to the inspection station, as dictated by the size and shape of forging
- Preparation of the forging is minimal, mainly involving the removal of surface contaminants that would prevent magnetization or inhibit particle mobility
- Routine inspection work can be effectively done by relatively unskilled labor properly trained in interpretation
- For forgings that are simple in configuration, and when justified by the quantity, magnetic particle inspection can be automated
- For some forgings, electronic sensing can be used, thus reducing the chances of human error and increasing inspection reliability
- Many forgings have sufficient retentivity to permit the use of multidirectional magnetization, thus permitting the inspection of indications in all orientations with a single preparation. Retentivity must be checked for the particular forging before a decision is made to use multidirectional magnetization

- The cost of magnetic particle inspection is generally lower than that for several other inspection methods in terms of investment in equipment, inspection materials, and inspection time

The limitations, generally the same as for inspecting other workpieces, of the magnetic particle inspection of forgings are:

- The method is applicable only to forgings made from ferromagnetic metals
- Because magnetic particle inspection is basically an aided visual inspection, under most circumstances, its effectiveness is subject to the visual acuity and judgment of the inspector
- Magnetic particle inspection is generally limited to detecting imperfections that are within about 3 mm (⅛ in.) of the surface of the forging
- Because the forging must be thoroughly magnetized, magnetic particle inspection is likely to be ineffective unless scale, grease, or other contaminants are removed from the forging. Such surface contaminants inhibit the mobility of the particles necessary to delineate the indications
- Following inspection, the forging usually must be demagnetized, depending mainly on the retentivity of the particular metal, subsequent shop operations, and end use

Detection of Surface Discontinuities. Magnetic particle inspection and liquid penetrant inspection are both widely used for detecting discontinuities in steel forgings, although the former is more widely used. As described previously, one advantage of using the magnetic particle technique is its ability to detect certain subsurface discontinuities that are not open to the surface. Subsurface discontinuities cannot be located with liquid penetrants. Also, some surface discontinuities may be so packed with scale that liquid penetrant techniques are marginal or infeasible. Therefore, in most cases, magnetic particle inspection is preferred to liquid penetrant inspection. Continuous magnetization is usually prescribed for inspecting steel forgings, because at the stage in which the forgings are inspected they are in an annealed or semiannealed condition and consequently have poor retentivity of magnetism.

Two inspection methods are available: dry powder and wet. Selection between the dry and the wet methods may sometimes be purely arbitrary, although it is usually based on the available equipment and the size of the forgings being inspected. The dry powder method is used to a greater extent for large forgings. Similarly, selection between fluorescent and nonfluorescent particles may often be arbitrary, although the size of the forging can be a major factor, because if the fluorescent method is used the forging must usually be of such size and shape that it can be inspected under ultraviolet light, with white light substantially eliminated.

Many specific procedures have been established for in-plant use. The dry powder and wet techniques adopted in one plant for the inspection of ferromagnetic metal forgings are described below.

Dry Powder Technique. The contact method of magnetization was selected to inspect the ferromagnetic materials. Prods are used to pass direct current or rectified alternating current through the workpiece. The magnetic particles are nontoxic, finely divided ferromagnetic material of high permeability and low retentivity, free from rust, grease, dirt, or other materials that may interfere with the proper functioning of the magnetic particles. The particles must also exhibit good visual contrast with the forging being inspected.

Inspection is by the continuous current method; that is, the magnetizing current remains on during the period of time that the magnetic particles are being applied and also while the excess particles are being removed. Prods are spaced 150 to 200 mm (6 to 8 in.) apart, except where restricted by configuration. The magnetic field is induced in two directions, 90° apart. The current used is 4 to 5 A/mm (100 to 125 A/in.) of prod spacing and is kept on for a minimum of 1/5 s.

Dry magnetic particles are applied uniformly to the surface, using a light dustlike technique. Excess particles are removed by a dry air current of sufficient force to remove the excess particles without disturbing particles that show indications.

The nozzle is held obliquely about 35 to 50 mm (1½ to 2 in.) above the test area. Nozzle size and air pressure result in a pressure (measured by a manometer) of 25 to 40 mm (1.0 to 1.5 in.) of water at an axial distance of 25 mm (1 in.) from the nozzle and 7.5 to 15 mm (0.3 to 0.6 in.) of water at 50 mm (2 in.) from the nozzle.

A 100 mm (4 in.) grid pattern over the entire forging surface is normally used for evaluation (Fig. 12). The prods are placed 200 mm (8 in.) apart, except where restricted by the shape of the forging, when using this grid pattern.

Prods are placed on the surface to be tested in the proper position, as shown by position 1 in Fig. 12, and the current is turned on (4 to 5 A/mm, or 100 to 125 A/in., of prod spacing). The powder is applied, the excess particles are removed, and the current is turned off. Inspection is conducted during application of the powder and after removal of the excess particles.

The next step is to reposition the prods 90°, as indicated by position 2 in Fig. 12; the procedure is then repeated. When the shape of the forging does not permit a full 90° rotation with the established prod spacing, the prod spacing can be changed, provided it is not less than 50 mm (2 in.) nor more than 200 mm (8 in.) between prods.

Wet Technique. The magnetic particles selected are nonfluorescent and suspended in a liquid vehicle. The magnetizing equipment is capable of inducing a magnetic flux of suitable intensity in the desired direction

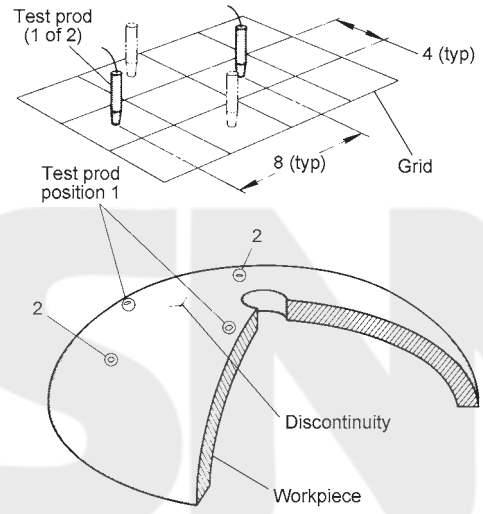


Fig. 12 Grid pattern and prod positions used in one plant for the magnetic particle inspection of forgings using prods and the dry powder technique. Dimensions are in inches. Source: Ref 1

by both the circular and the longitudinal methods. Direct current from generators, storage batteries, or rectifiers is used to induce the magnetic flux.

Circular magnetization is obtained by passing the current through the forging being examined or through a central conductor to induce the magnetic flux. Longitudinal magnetization is obtained by using a solenoid, coil, or magnet to induce the magnetic flux. The magnetic particles are nontoxic and exhibit good visual contrast. The viscosity of the suspension vehicle for the particles must be a maximum of 5×10^{-6} m^2/s (5.0 centistokes) at any bath temperature used. The magnetic particles are limited to 28 to 40 g (1.0 to 1.4 oz) of solid per gallon of liquid vehicle. The liquid used as a vehicle for the magnetic particles may be a petroleum distillate such as kerosene. Tap water with suitable rust inhibitors and wetting and antifoaming agents can be substituted for the petroleum distillate. The water should contain about 0.3% antifoam agent, 3.9% rust inhibitor, and 12.8% wetting agent.

Inspection is carried out by the continuous method. For this method, the magnetizing circuit is closed just before applying the suspension or just before removing the forging from the suspension. The circuit remains closed for approximately a half second.

For circular magnetization, an ammeter is used to verify the presence of adequate field strength. For verifying adequate field strength in longitudinal magnetization, a field indicator is useful. Typical current levels utilized to provide an adequate field strength are 4 to 12 A/mm (100 to 300

A/in.) of diameter of the surface being examined, although current levels of up to 30 A/mm (750 A/in.) of diameter have been used.

The magnetizing force for longitudinal magnetization is 2000 to 4000 ampere-turns per 25 mm (1 in.) of diameter of the surface being examined. If both the inside and outside diameters of cylindrical parts are to be inspected, the larger diameter is used in establishing the current. If it is impractical to attain currents of the calculated magnitude, a magnetic field indicator is used to verify the adequacy of the magnetic field.

Suspensions must be tested daily or when they appear to have become discolored by oil or contaminated by lint. Common practice is to test the suspension at the beginning of each operating shift. Steps of the suspension test are:

1. Let the pump motor run for several minutes to agitate a normal mixture of particles and vehicle
2. Flow the bath mixture through the hose and nozzle for a few minutes to clear the hose
3. Fill the centrifuge tube to the 100 mL line
4. Place the centrifuge tube and stand in a location free from vibration
5. Let the tube stand for 30 minutes for particles to settle out
6. After 30 minutes, readings for settled particles should be 1.7 to 2.4 mL. If the reading is higher, add vehicle; if lower, add particle powder to the suspension

Liquid Penetrant Inspection

Liquid penetrant inspection is widely used for locating surface imperfections in all types of forgings, either ferrous or nonferrous, although it is more frequently used on nonferrous forgings. There is no limitation on the size or shape of a forging that can be liquid penetrant inspected.

Any of the three basic liquid penetrant systems (water-washable, postemulsifiable, and solvent-removable) can be used to inspect forgings. The product or product form is not a principal factor in the selection of a system.

Advantages. Among the advantages of liquid penetrant inspection of forgings are:

- There are no limitations on metal composition or heat treated condition
- There are no limitations imposed on the size or shape of the forging that can be inspected
- Liquid penetrant inspection can be done with relatively simple equipment
- Training requirements for inspectors are minimal
- Inspection can be performed at any stage of manufacture

- Liquid penetrant materials can be taken to the forgings or the forgings taken to the inspection station, depending on the size and shape of the forgings

The limitations of the liquid penetrant inspection of forgings are basically the same as those for the inspection of other workpieces. The characteristics of the surface of a forging sometimes impose specific limitations. The most important general limitations are:

- Liquid penetrant inspection is restricted to detecting discontinuities that are open to the surface
- Liquid penetrant inspection is basically a visual aid; therefore, results depend greatly on the visual acuity and judgment of the inspector
- Satisfactory inspection results require that the surface of the forging be thoroughly cleaned before inspection. The presence of surface scale can cause inaccurate readouts. If the surface of the forging is excessively scaled, it should be pickled or grit blasted, preferably pickled. The forgings should also be cleaned to remove surface contaminants, such as grease and oil
- Liquid penetrant inspection is slower than magnetic particle inspection

Liquid Penetrant Detection of Flaws in Steel Forgings

Factors affecting the selection of a special penetrant system for inspecting steel forgings include available equipment; size, shape, and surface conditions of the forgings; degree of sensitivity required; whether or not the entire forging requires inspection; and cost. Regardless of which system is used, the degree of success achieved depends greatly on the surface conditions of the forging. Rough, scaly surfaces are likely to result in either false indications or obscuring of meaningful flaws. Pickled surfaces are preferred. Abrasive blasting is usually satisfactory for cleaning forging surfaces, although overblasting must be avoided or some flaws may be tightly closed preventing the penetrant from entering. The postemulsifiable and solvent-removable liquid penetrant systems are most often used to inspect steel forgings.

The postemulsifiable system is generally preferred to the water-washable system for forgings because of its greater sensitivity. Either the fluorescent-penetrant or the visible-dye technique can be used. Selection depends largely on whether ultraviolet light inspection can be used. For forgings of a size and shape that can be immersed in tanks and inspected in a booth, the fluorescent technique is usually preferred.

The solvent-removable system is especially well adapted to applications in which only a portion of the forging requires inspection. Equipment for this system can be minimal and completely portable or may in-

volve more elaborate systems used on a production basis, as described in the following examples.

Liquid Penetrant Detection of Flaws in Heat-Resistant Alloy Forgings

Because most heat-resistant alloy forgings are nonmagnetic, the use of magnetic particle inspection for detecting surface flaws cannot be considered. Liquid penetrants are extensively used for inspecting the surfaces of high integrity forgings.

Critical forgings such as these require close quality control surveillance. Following production penetrant inspection using group VI fluorescent penetrant, it is desirable to conduct quality control overchecks on samples selected from previously inspected batches. These overchecks are performed using highly sensitive, high resolution penetrants.

Ultrasonic Inspection

Ultrasonic inspection is used to detect both large and small internal flaws in forgings. Forgings, by their nature, are amenable to ultrasonic inspection. Both longitudinal or shear wave (straight or angle beam) techniques are utilized. The size, orientation, location, and distribution of flaws influence the selection of technique and the inspection results.

Consider, for example, Fig. 13, which shows the influence of flaw orientation on signal response.

There are, however, some definite limitations. All ultrasonic systems generate sound electrically and transmit the energy through a transducer to the forging. Because the relationship of sound transmitted to sound re-

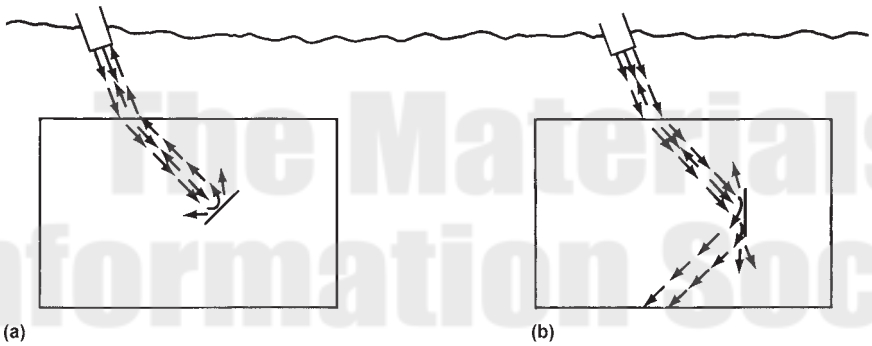


Fig. 13 Comparisons of discontinuities at normal orientation versus radial orientation. (a) Discontinuity normal to sound energy, almost total reflection of energy back toward transducer. (b) Discontinuity at 0° with respect to 45° sound energy path, almost no energy reflected back toward transducer. Source: Ref 1

ceived is a factor in the inspectability of a forging, particular attention must be given to the surface condition of the forging.

Although techniques and couplants can enhance the energy transmission from the transducer to the forging, as-forged surfaces impair the effectiveness of ultrasonic inspection. Near-surface flaws are most difficult to detect, and a dead zone at the entry surface often interferes. Because of the difficulty involved in detecting surface flaws by ultrasonic inspection, another method, such as magnetic particle or liquid penetrant inspection, is often used in conjunction with ultrasonic inspection to inspect high integrity forgings thoroughly.

Complex shapes are difficult to inspect ultrasonically because of the problems associated with sound entry angle. Most ultrasonic inspection of forgings uses techniques that send waves into the forging perpendicular to the surface. Radii, fillets, and similar configurations must receive special treatment if all areas of the forging must be inspected. The special treatment involves the use of a standoff that has an end contoured to fit the inspection surface or the use of a small diameter or focused transducer.

In certain cases, where the end use of a forging is considered critical, ultrasonic inspection is used to inspect the wrought material before it is worked. Surface or internal flaws that are not detected before a billet is forged may not be detected in the final forging and will be present in the finished part. Ultrasonic inspection is often used as part of a completely diagnostic inspection of a forging from newly designed dies, where use of the finished part does not warrant inspection of every part. Quality control measures often include the ultrasonic inspection of random samples from a particular forging. This provides the necessary assurance that the process is under control and that variables affecting internal quality have not been inadvertently introduced. Ultrasonic inspection is often used in the further evaluation of flaws detected by other nondestructive methods. This reduces the possibility that a particular forging will be unsuitable for its intended service.

Ultrasonic inspection can be used on every forging to validate its integrity for extremely rigorous requirements. This applies in particular to forgings for nuclear and aerospace applications, where rigid standards of acceptance have been established. Standards and criteria have been set up to detect material inclusions, internal voids, laminations, and other conditions. In addition, the inspection of every forging by ultrasonics has been effective in detecting excessive grain size and other structural conditions.

Ultrasonic inspection is often used to qualify a particular lot of forgings that has been subjected to certain variations in approved processing procedures. A notable instance is the use of ultrasonics to determine the presence of thermal flakes or in locating quench cracks.

Radiographic Inspection

Radiography (γ-ray or x-ray) is not extensively used for the inspection of forgings for two reasons. First, the types of discontinuities most commonly located by radiography (gas porosity, shrinkage porosity, and shrinkage cavities) are not usually found in forgings. Second, for the types of internal discontinuities that are commonly found in forgings (inclusions, pipe, bursts, or flakes), ultrasonic inspection is more effective, more adaptable, and more economical.

Radiographic techniques can sometimes be helpful in the further investigation of known internal discontinuities in forgings when the presence of these discontinuities has been determined earlier by ultrasonic inspection. In sections that are not too thick to penetrate with available radiographic equipment, the size, orientation, and possibly the type of discontinuities can be evaluated by radiography.

ACKNOWLEDGMENT

This chapter was adapted from Nondestructive Inspection of Forgings, *Nondestructive Evaluation and Quality Control*, Volume 17, *ASM Handbook*, 1992.

REFERENCES

1. Nondestructive Inspection of Forgings, *Nondestructive Evaluation and Quality Control*, Vol 17, *ASM Handbook*, ASM International, 1992, p 491–511

CHAPTER 16

Inspection of Powder Metallurgy Parts

FABRICATED POWDER METALLURGY (P/M) parts are evaluated and tested at several stages during manufacturing for part acceptance and process control. The various types of tests included are:

- Dimensional evaluation
- Density measurements
- Hardness testing
- Mechanical testing
- NDT

Dimensional Evaluation

Dimensional accuracy of P/M sintered parts is determined with the same measurement techniques that are used for wrought materials. However, other testing methods for P/M materials are specialized, such as determination of surface finish. For sintered parts, a chisel pointed stylus is used to deemphasize the effects of porosity. A conical stylus tends to engage porosity, thus giving an exaggerated measurement of roughness.

During the manufacture of sintered parts, dimensional change must be accommodated for during each processing step. Causes of dimensional changes include:

- Elastic springback during ejection from the tooling used for cold pressing
- Growth or shrinkage during delubrication, presintering, and sintering
- Elastic springback from tooling during cold repressing or sizing

- Thermal contraction from the tools used in hot forging or hot repressing
- Tool wear in cold or hot compacting
- Machining tolerances at secondary machining and associated tool wear
- Distortion during annealing
- Growth or shrinkage during carburizing, nitriding, or neutral hardening
- Shrinkage during tempering
- Growth during steam blackening

Parts manufacturers must be familiar with the amount of dimensional changes to expect for the materials and equipment in use so that tooling can be produced that accommodates these changes and produces accurate parts. Understanding and controlling these factors is essential to commercial P/M parts manufacturing.

Density Measurement

Density is the ratio of mass to volume. For a given material, degree of sintering, and heat treatment, density determines mechanical and physical properties. For example, higher density in sintered steels results in higher tensile strength, elongation, and impact resistance. As-pressed or green density also influences growth or shrinkage that occurs during sintering. With nonuniform green density, parts grow or shrink nonuniformly, as in a thin walled bronze bearing with a low density region equidistant from the ends. This results in a significantly smaller diameter at mid-length than at the ends and necessitates repressing or sizing for close dimensional control.

If cubes or right cylinders can be extracted from actual parts, linear dimensions can be measured and volume can be easily calculated. From the weight of a part, density can be calculated. This yields a value that, under ideal conditions, differs by 0.04 g/cm^3 (0.5%) from a reference. Unless the sintered part is directly molded to an easily measured shape such as a transverse rupture bar (31.8 by 12.7 by 6.4 mm or 1.25 by 0.50 by 0.25 in.), this method of measuring linear dimensions is infrequently used.

Methods Based on Archimedes' Principle. Typical methods of measuring density depend on Archimedes' principle, in which hydrostatic forces in a liquid exerts buoyant forces proportional to the part volume. This measurement is standardized in ASTM B328, MPIF test method 42, and International Standards Organization test method ISO 2738.

When an object is immersed in a liquid, the liquid exerts an upward buoyant force that is equal to the product of the object volume and the density of the liquid. The difference in weight between an object weighed

in air and its weight when suspended in water is equal to the object volume in cubic centimeters times the density of water. Approximating the density of water as unity:

$$V = W_{air} - W_{water}$$

where V is the volume in cm^3, W_{air} is the weight in air in g, and W_{water} is the weight of object suspended in water less the weight of the suspending wire in water (tare) in g. Density in g/cm^3 is then:

$$\text{Density} = W_{air}/(W_{air} - W_{water})$$

For unsintered materials molded with 0.75% lubricant, the pores are well sealed, and water cannot penetrate. For such parts, the above calculation is suitable. It is also suitable for materials with pores that are sealed off from the surface (materials close to theoretical density).

However, for most sintered materials that are 70 to 95% dense, water tends to infiltrate the pores during weighing in water. This minimizes the buoyancy effect of the water (that is, the liquid is acting on a smaller volume) and results in an erroneous calculation of low volume. This low volume then causes an erroneously high density value. Infiltration of water into pores usually is accompanied by air bubbles escaping from the part. If the part is blotted to remove surface water and reweighed in air after weighing in water, any weight gain indicates that water has entered the pores. Although not a standard procedure, volume can be approximated as the weight in air after removing the part from the water, minus the weight in water.

To prevent infiltration of water, all three standard test methods require that the pores of the part be filled with oil. Oil impregnation is done after the part is weighed in air. This is carried out under vacuum or by immersion in hot oil. Oil prevents the water from entering the pores. The volume of the part is then determined as the part weight in air with oil in the pores, minus the weight of the oiled part suspended in water. Care should be taken to select an oil that is not soluble in water or not soluble in water plus wetting agent. Such oils also must exhibit superior demulsibility.

The precision of the ISO method is ±0.25%, regardless of sample density, and assumes a water density of 0.997 g/cm^3. To determine density variation from one point to another in a complex part, the available samples must be considerably <17 g. According to ASTM B328, a minimum sample of 2 g is recommended, because a relatively high error rate results from measuring small samples.

Metallographic estimates can also be made of the area fraction of porosity, which is numerically equal to the volume porosity, and thus the density of sintered materials. The method is not standardized, and accu-

racy of results depends on the skill of the metallographer to define the correct area fraction of porosity. Frequently, the amount of porosity is exaggerated or minimized because of over polishing or under polishing. Prior to mounting, it is recommended that any oil and cutting fluids be removed from the pores by Soxhlet extraction or heating in air, followed by impregnating the pores with epoxy resin. This method can achieve a precision of about ±0.1 g/cm^3, depending on the laboratory. The advent of metallography with a television monitor and quantitative metallographic functions allows rapid measurement of area fractions of pores. However, this method is highly dependent on proper sample preparation.

Apparent Hardness and Microhardness

Porous materials exhibit a wider variation in hardness testing than their wrought counterparts. The entrance of the indenter into pores or groups of pores generally causes this effect. At least five consistent readings should be taken, in addition to any obviously high or low readings, which are discarded. The remaining five readings should be averaged.

Because most published data show typical hardness values, the buyer and seller must agree on specified or minimum values. The seller and user of P/M materials also should agree on which area or areas of a part are to be hardness tested. The average of five or more consistent readings must meet the standard hardness, not any single reading. Recommended scales for taking accurate measurements are summarized in Table 1.

Microhardness of porous materials can best be measured with Knoop or diamond pyramid hardness indenters at loads of 100 g (0.2 lb) or greater. In atomized irons, particles exhibit minimal porosity; consequently, the Knoop indenter is suitable. It makes a very shallow indentation and is only infrequently disturbed by entering undisclosed pores. Care should be exercised in preparing the sample surface. Use of the diamond pyramid indenter is particularly well suited to irons, which contain numerous fine internal pores. Because of its greater depth of penetration, the diamond pyramid indenter frequently encounters hidden pores. Microhardness testing and measurement of case depth are covered by Metal Powder Industries Federation standard MPIF 37.

Table 1 Common hardness scales

Material	Sintered hardness scale	Heat treated hardness scale
Iron	HRH, HRB	HRB, HRC
Iron-carbon	HRB	HRB, HRC
Iron-nickel-carbon	HRB	HRC
Prealloyed steel	HRB	HRC
Bronze	HRH	…
Brass	HRH	…

Scource: Ref 1

Mechanical Testing/Tensile Testing

Metal Powder Industries Federation standard 10 describes specimens for tensile testing. Specimens include flat unmachined test bars or machined rounds. Unmachined flats are more prone to slippage of grips or breakage in the gauge region. For testing to be meaningful, it is important to verify that such bars are free of microlaminations, which requires careful metallographic evaluation. Highest quality bars are molded in well bolstered die sets (890 kN or 100 ton rating), with generous exit taper, high green strength powder, and top punch hold down, if possible.

With all steel test bars, it is necessary to determine that substantial carburization or decarburization resulting from sintering or hardening does not exist. If the bars are heat treated, the microstructure at the surface and in the interior should be described in the test report, because many P/M steels have low hardenability.

For heat treated materials, unmachined flat specimens tend to slip in the grips. A machined bar provides more accurate data. The machined bar shows an increase in apparent tensile strength of 50% compared to molded bars. Even with machined bars, some heat treated materials exhibit such low elongation that failure occurs prior to reaching 0.2% permanent deformation.

Transverse Rupture Strength. The transverse rupture test breaks a 31.8 by 12.7 by 6.4 mm (1.25 by 0.5 by 0.25 in.) test bar as a simple beam. The test is theoretically valid only for perfectly brittle materials and measures the stress in the outer fiber at fracture. For many sintered steels, transverse rupture stress is considered to be equal to twice the ultimate tensile strength. This is only an approximation. Metal Powder Industries Federation standard 35 presents transverse rupture stress values that correspond with ultimate tensile stress values for common P/M materials.

The transverse rupture test is useful for comparing and evaluating materials for strength, even if the bar bends before fracture. These bars are preferred, because they can be molded and sintered conveniently. Testing is faster than when using a tensile bar. When heat treated, the test bar does not experience distortion. This testing procedure is used mainly as a quality control tool to ensure the maintenance of minimum mechanical properties. During P/M part production, this method is used to evaluate properties of incoming powder, such as compressibility, sintered strength, and dimensional change.

Unnotched Charpy Impact Strength. A 10 by 10 by 76 mm (0.35 by 0.35 by 3.00 in.) molded bar is used for impact testing. The unnotched bar provides a more sensitive test, suitable for use on materials with an impact strength below 14 J (10 ft · lbf). The bar is conventionally struck on the surface that contacted the die at molding.

Proof Testing. The most common method of demonstrating the strength of sintered parts is through mechanical tests that stress parts to failure. Qualification samples or first production lots are used to establish desired strength values; these data are incorporated in the part design specification.

For testing gears, several teeth are removed. The remaining teeth are loaded in a predetermined arrangement on a fixture. The load to fracture is recorded. To be meaningful, the destructive test must imitate service loading on the part.

Impact or drop weight tests also are used to evaluate materials. A drop weight test does not only use an acceptance or rejection evaluation. This testing procedure investigates the impacts above the acceptance level and below the rejection value. When a part does not break, each succeeding load is increased until breakage occurs. Thus, if load *P* does not break a part, 1.05 *P* is used on the succeeding impact.

Powder Metallurgy Part Defects

The problem of forming defects in green parts during compaction and ejection has become more prevalent as parts producers have begun to use higher compaction pressures in an effort to achieve high density, high performance powder metallurgy (P/M) steels. Proper press setup for molding P/M parts is also critical to prevent cracking. In a flanged part that experiences a change in diameter, density in the hub and flange should be nearly equal. Unequal density leads to powder displacement from one part level to the next and to the formation of shear cracks. Such cracks often occur at 45° to the pressing direction and at surfaces at the junction (radius) between the hub and flange. At press setup, equal density should be obtained in the hub and flange. A high green strength powder and a press that maintains a small counter pressure on the top of the part during ejection from the tools (top punch hold-down) should be used.

The four most common types of defects in P/M parts are ejection cracks, density variations, microlaminations, and poor sintering.

Ejection Cracks. When a part has been pressed, there are large residual stresses in the part due to the constraint of the die and punches, which are relieved as the part is ejected from the die. The strains associated with this stress relief depend on the compacting pressure, the green expansion of the material being compacted, and the rigidity of the die. Green expansion, also known as *spring out*, is the difference between the ejected part size and the die size. A typical value of green expansion for a powder mix based on atomized iron powder pressed at relatively high pressure (600 to 700 MPa or 45 to 50 tsi) is 0.20%. For example, in a partially ejected compact, the portion that is out of the die expands to relieve the residual stresses, while the constrained portion remains die size and a shear stress is imposed on the compact. When the ability of the powder compact to

accommodate the shear stress is exceeded, ejection cracks such as the one shown in Fig. 1 are formed.

The radial strain can be alleviated to a degree by increasing the die rigidity and designing some release into the die cavity. However, assuming that the ejection punch motions are properly coordinated, the successful ejection of multilevel parts depends to a large degree on the use of a high quality powder that combines high green strength with low green expansion and low stripping pressure.

Density Variations. Even in the simplest geometry possible, a solid circular cylinder, conventional pressing of a part to an overall relative density of 80% will result in a distribution of density within the part ranging from about 72 to 82%. The addition of simple features, such as a central hole and gear teeth, presents minor problems compared to the introduction of a step or second level in the part. Depending on the severity of the step, a separate, independently actuated punch can be required for each level of the part. During the very early stages of compaction, the powder redistributes itself by flowing between sections of the die cavity. However, when the pressure increases and the powder movement is restricted, shearing of the compact in planes parallel to the punch axis can only be avoided by proper coordination of punch motions. When such shear exists, a density gradient results and is not always severe enough for an associated crack to form upon ejection. However, a low density area around an internal corner, as shown in Fig. 2, can be a fatal flaw, because this corner is usually a point of stress concentration when the part is loaded in service.

Microlaminations. In photomicrographs of unetched part cross sections, microlaminations such as those shown in Fig. 3 appear as layers of unsintered interparticle boundaries that are oriented in planes normal to the punch axis. They can be the result of fine microcracks associated with shear stresses upon ejection; such microcracks fail to heal during sintering. Because of their orientation parallel to the tensile axis of standard

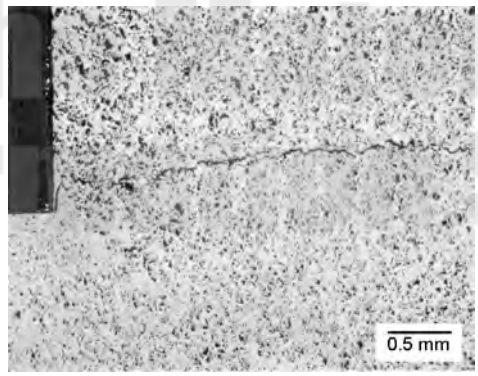


Fig. 1 Ejection crack in sintered P/M steel, unetched. Source: Ref 2

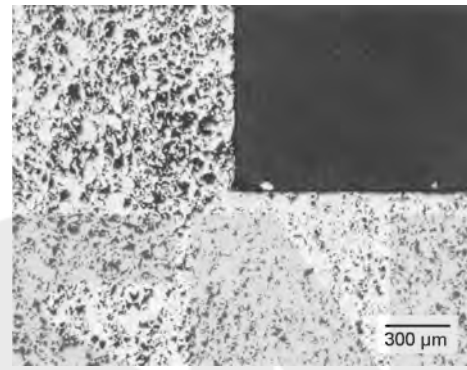


Fig. 2 Density gradient around an internal corner in a part made with a single piece stepped punch, unetched. Source: Ref 1

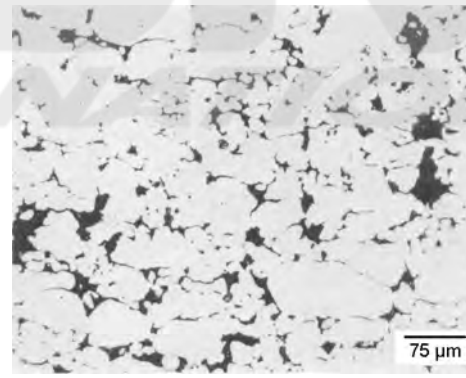


Fig. 3 Microlaminations in sintered P/M steel, unetched. Source: Ref 1

test bars, they have little influence on the measured tensile properties of the bars, but are presumed to be a cause of severe anisotropy of tensile properties.

Poor Sintering. When unsintered particle boundaries result from a cause other than shear stresses, they are usually present because of insufficient sintering time or sintering temperature, a nonreducing atmosphere, poor lubricant burn-off, inhibition of graphite dissolution, or a combination of these. A severe example is shown in Fig. 4. Unlike microlaminations, defects associated with a poor degree of sintering are not oriented in planes.

Flaw Detection

Crack detection is accomplished by various methods, such as mechanical proof testing, metallography, and filtered particle or magnetic particle inspection. Other nondestructive testing methods for P/M applications also include electrical resistivity testing, eddy current, and magnetic bridge

testing, magnetic particle inspection, ultrasonic testing, x-ray radiography, gas permeability testing, and gamma-ray (γ-ray) density determination. The capabilities and limitations of these techniques are summarized in Table 2.

Mechanical Proof Testing. A sampling of sintered parts can be broken to confirm the presence of a suspected cracking problem. For example, the

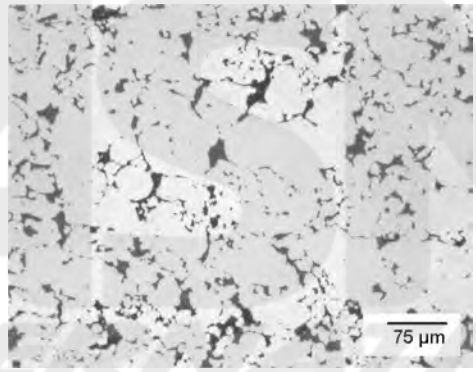


Fig. 4 Poor degree of sintering in P/M compact, unetched. Source: Ref 2

Table 2 Comparison of the applicability of various nondestructive evaluation methods to flaw detection in P/M parts

Method	Measured/detected	Applicability to PM parts(a)		Advantages	Disadvantages
		Green	Sintered		
X-ray radiography	Density variations, cracks, inclusions	C	C	Can be automated	Relatively high initial cost; radiation hazard
Computed tomography	Density variations, cracks, inclusions	C	C	Can be automated; pinpoint defect location	Extremely high initial cost; highly trained operator required; radiation hazard
Gamma-ray density determination	Density variations	A	A	High resolution and accuracy; relatively fast	High initial cost; radiation hazard
Ultrasonic imaging, C-scan	Density variations, cracks	D	B	Sensitive to cracks; fast	Coupling agent required
Ultrasonic imaging, SLAM	Density variations, cracks	D	C	Fast; high resolution	High initial cost; coupling agent required
Resonance testing	Overall density, cracks	D	B	Low cost; fast	Does not give information on defect location
Acoustic emission	Cracking during pressing and ejection	C	D	Low cost	Exploratory
Thermal wave imaging	Subsurface cracks, density variations	D	C	No coupling agent required	Flat or convex surfaces only
Electrical resistivity	Subsurface cracks, density variations, degree of sinter	A	A	Low cost, portable, high potential for use on green compacts	Sensitive to edge effects
Eddy current/magnetic bridge	Cracks, overall density, hardness, chemistry	C	A	Low cost, fast, can be automated; used on P/M valve seat inserts	Under development
Magnetic particle inspection	Surface and near-surface cracks	C	A	Simple to operate, low cost	Slow; operator sensitive
Liquid dye penetrant inspection	Surface cracks	C	D	Low cost	Very slow; cracks must intersect surface
Pore pressure rupture/gas permeability	Laminations, ejections, cracks, sintered density variations	A	A	Low cost, simple fast	Gas-tight fixture required; cracks in green parts must intersect surface

(a) A, has been used in the production of commercial P/M parts; B, under development for use in P/M; C, could be developed for use in P/M, but no published trials yet; D, low probability of successful application to P/M. Source: Ref 2

flanges can be pushed off hubs under shear at the suspected crack location. The presence of a few unexplained low readings indicates that an initiating crack is present.

Metallography. Low powered binocular microscopes can be used to detect cracks at changes in diameter. Metallography is a more time-consuming method. A sampling of parts is sectioned parallel to the pressing direction. When mounted and carefully polished to expose open pores and cracks, the presence of minute cracks is apparent (Fig. 5).

Liquid Penetrant Crack Detection. Most sintered parts have porous surfaces that absorb and then release sufficient penetrant in all regions so that it is impossible to distinguish the crack from the porosity background. The dye does not preferentially reside at cracks in P/M parts, because the pore radius and the crack radius are equivalent.

Filtered Particle Crack Detection. One proprietary process of filtered particle crack detection (Partek) involves brief immersion of the test piece in a suspension of fluorescent particles. Particles are filtered and collect near the surface of cracks as the fluid enters. Cracks are clearly visible under black light. This one step method is used to detect cracks in presintered porous tungsten carbide blanks. To the extent that an unsintered part has open porosity, this method also can be used on green parts. Density cannot be too high, however, and excessive lubricant tends to clog the pores. Successful use of this method on presintered porous tungsten carbide blanks indicates that it may be suitable for sintered P/M parts with open pores into which fluid can enter. Small cracks fluoresce brightly, while large cracks are darker than the surrounding fluorescing surface.

Magnetic particle crack inspection can detect some near-surface cracks in sintered parts. However, unsintered parts are not adequately bonded to support a magnetic flux, and this method is consequently unsatisfactory. Magnetic particle detection methods have been successfully used for many years for inspecting medium density sintered automotive parts, by both P/M parts producers and automotive manufacturers. It is possible to automate the inspection process by using digital image processing.

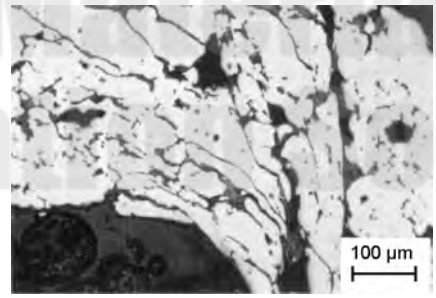


Fig. 5 Cracks and unbounded particles at the junction (radius) of a horizontal flange and vertical hub resulting from shear during compaction, unetched. Source: Ref 1

Direct Current Resistivity Testing. A voltage field within a conductive solid will create currents that are influenced by structural irregularities, including cracks and porosity. The arrangement shown in Fig. 6 is used to measure the voltage drop in a current field localized between two electrode probes. This method has been used to detect seeded defects in laboratory specimens and has also been successfully applied to the production of sintered steel parts.

The direct current resistivity test can be used on any conductive material; it is not limited to ferromagnetic materials. Although further development is needed, resistivity measurements appear to be a promising technique for the nondestructive evaluation of both green and sintered P/M parts. In addition to detecting cracks in green parts, as well as part-to-part density variation, studies have shown that changes in resistivity due to poor carbon pickup during sintering were also detectable. Resistivity testing has also been used later in the processing sequence to screen heat treated parts for incomplete transformation to martensite. Another study has yielded the relative density/conductivity relationship, suggesting that resistivity tests could be used as a rapid check for localized density variations. As with ultrasound, the elastic modulus and the toughness of porous steels can also be distinguished by resistivity checks.

Radiographic Techniques

X-Ray Radiography. Any feature of a part that either reduces or increases x-ray attenuation will be resolvable by x-ray radiography. Some types of flaws and their x-ray images are shown in Fig. 7. The ability to detect defects depends on their orientation to the x-ray source. A crack parallel to the x-rays will result in reduced attenuation of the rays, and the x-ray film will be darker in this region. A thin crack perpendicular to the x-ray will hardly influence attenuation and will not be detected.

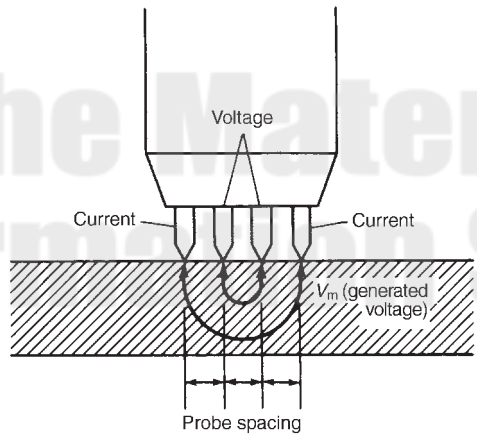


Fig. 6 Four-point probe used in the resistivity test. The outer probe pins are the current leads; the inner pins are the potential leads. Source: Ref 3

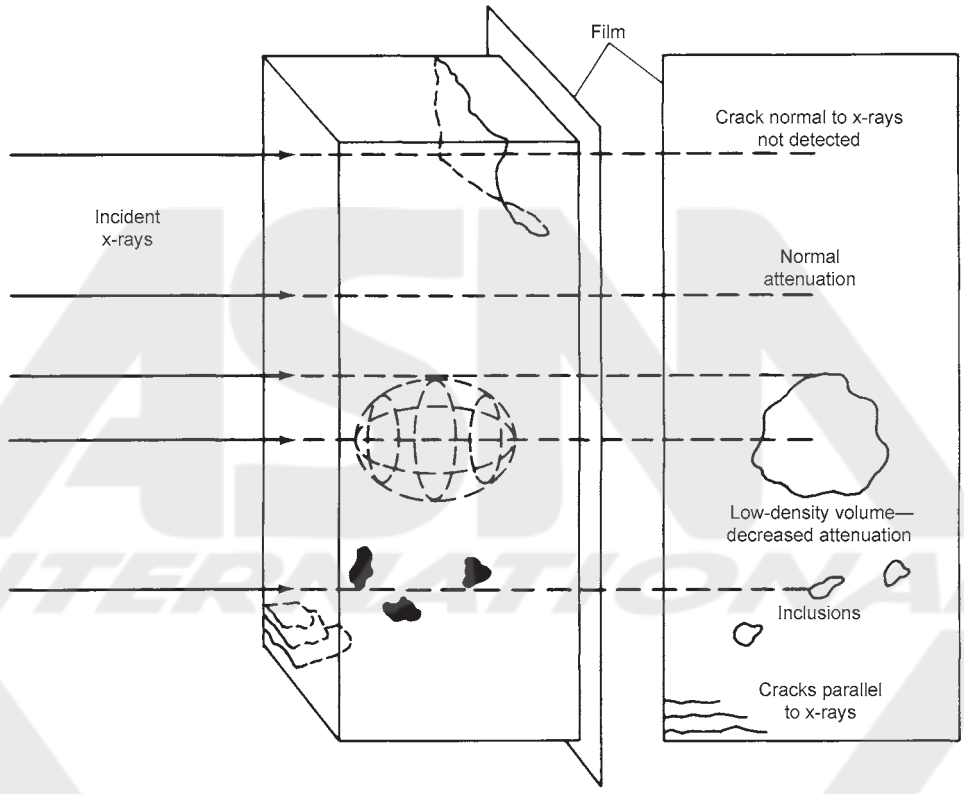


Fig. 7 Schematic of flaws and their x-ray images. Defect types that can be detected by x-ray radiography are those that change the attenuation of the transmitted x-rays. Source: Ref 4

Historically, flaw detection by x-ray radiography has been an expensive and cumbersome process suited only to safety critical and high value added parts. The process has been considerably improved by the development of real-time imaging techniques that replace photographic film. Real-time imaging allows parts to be tested rapidly and accepted or rejected on the spot.

Computed tomography is a version of x-ray radiography that includes highly sophisticated analysis of the detected radiation. A tomographic setup consists of a high energy photon source, a rotation table for the specimen, a detector array, and the associated data analysis and display equipment, as illustrated in Fig. 8. The ability to rotate the specimen increases the chance of orienting a defect relative to the x-rays such that it will be detected. The x-ray source and detector array can be raised or lowered to examine different planes through the sample.

In a typical system, the photon source can be a radioisotope such as ^{60}Co, depending on the energy requirements of the individual specimens. The lead aperture around the source acts as a collimator to produce a fan

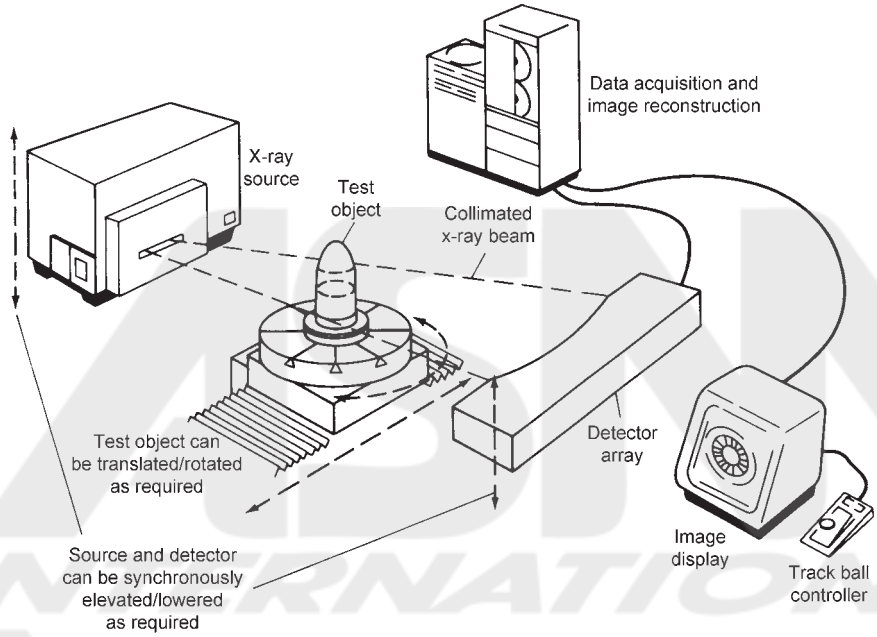


Fig. 8 Schematic of computed tomography, which is the reconstruction by computer of a series of tomographic planes (slices) of an object. The transmitted intensity of the fan shaped beam is processed by computer and the resulting image is displayed on a terminal. Source: Ref 5

shaped beam about 5 mm (0.2 in.) thick. The sample is rotated in incremental steps, and the transmitted radiation is detected at each step by computer controlled detectors situated one every 14 mm (0.55 in.) in a two-dimensional array.

The computer then reconstructs the object using intensity data from a number of scans at different orientations. The output is in the form of a two-dimensional plan in which colors are mapped onto the image according to the intensity of the transmitted radiation. The resolution available depends on the difference in density between the various features of the object. Experiments with P/M samples have shown that density can be measured to better than 1% accuracy, with a spatial resolution of 1 mm (0.04 in.).

Gamma-Ray Density Determination. Local variations in the density of P/M parts have been detected by measuring the attenuation of γ-rays passing through the part. Depending on the material and the dimensions of the part, density can be measured to an accuracy of ±0.2 to ±0.7%. The technique has been used by P/M parts fabricators in place of immersion density tests as an aid in tool setting.

The apparatus consists of a vertically collimated γ-ray beam originating from a radioisotope. The beam passes through the sample, as shown in

Fig. 9, and reaches a detector via a 1 mm (0.04 in.) diameter aperture, where the transmitted intensity is measured. The detector consists of a sodium iodide scintillation crystal, which in turn excites a photomultiplier. Exposure time is 1 to 2 minutes; a 4 mm (0.15 in.) aperture can reduce this time to 30 seconds at the expense of some resolution. The radiation source of the Gamma Densomat is Americium 241 (60 keV). For high energy beams, Cesium 137 (660 keV) can be substituted.

This method has been shown to be particularly useful in cases where the section of the part to be checked is too small for immersion density measurements

Ultrasonics

Many characteristics of solids can be determined from the behavior of sound waves propagating in them. Ultrasonic signals impinged on a sample at one surface are transmitted at speeds and attenuated at rates determined by the density, modulus of elasticity, and continuity of the material.

Green Compacts. The characterization of green compacts by ultrasonic techniques appears to be hindered by problems of extreme attenuation of the incident signal. In one case, signals of 1 to 20 MHz were transmitted through an 8 mm (0.3 in.) thick compact of atomized iron with 0.2% graphite added. Attenuation did not allow back-wall echo measurement. Density was found to influence the transmitted intensity, with specimens at 95% relative density allowing some degree of transmission over the entire range of frequencies tested, while specimens at 87% relative density damped the incident signals entirely. It has also been shown that the velocity of ultrasound in green parts is highly anisotropic and that the experimental reproducibility is very poor.

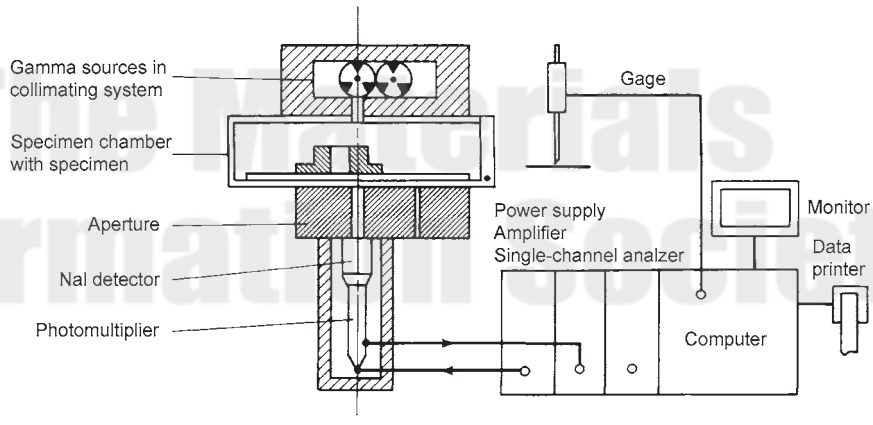


Fig. 9 Schematic of the Gamma Densomat. Source: Ref 6

Sintered Parts. In sintered parts, both the velocity of sound and their resonant frequencies have been related to density, yield strength, and tensile strength. Plain carbon steel P/M specimens were used in one series of tests, and the correlation was found to be close enough for the test to be used as a quick check for the degree of sintering in production P/M parts. Other work has demonstrated the relationship between sound velocity and tensile strength in porous parts (Fig. 10). The same types of relationships have also been documented in powder forgings.

Sintered parts have been found to transmit ultrasound according to the relationships shown in Fig. 11. The highest wave velocities occurred in the pressing direction. An additional distinction was found between the velocities in the longitudinal and lateral axes of an oblong specimen, and these results were shown to be reproducible between different powder lots and specimen groups. The anisotropy of velocity diminished at higher densities and disappeared above 6.85 g/cm^3.

The relationship between ultrasonic velocity and tensile yield and ultimate tensile strength is shown in Fig. 12 for as-sintered 0.65% carbon steel. Sintering times and densities that resulted in various tensile properties and velocities are given in Table 3. With similar data for a specific part, ultrasonic velocity can be easily measured, thereby determining the state of sintering and mechanical properties.

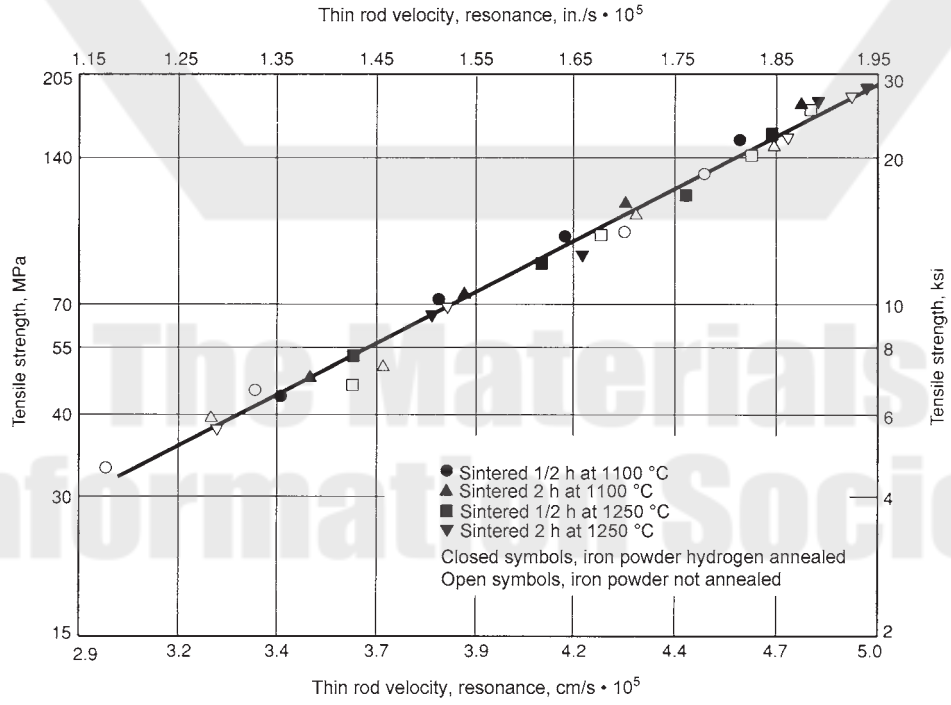


Fig. 10 Correlation of ultrasonic velocity with tensile strength of sintered steel. Source: Ref 7

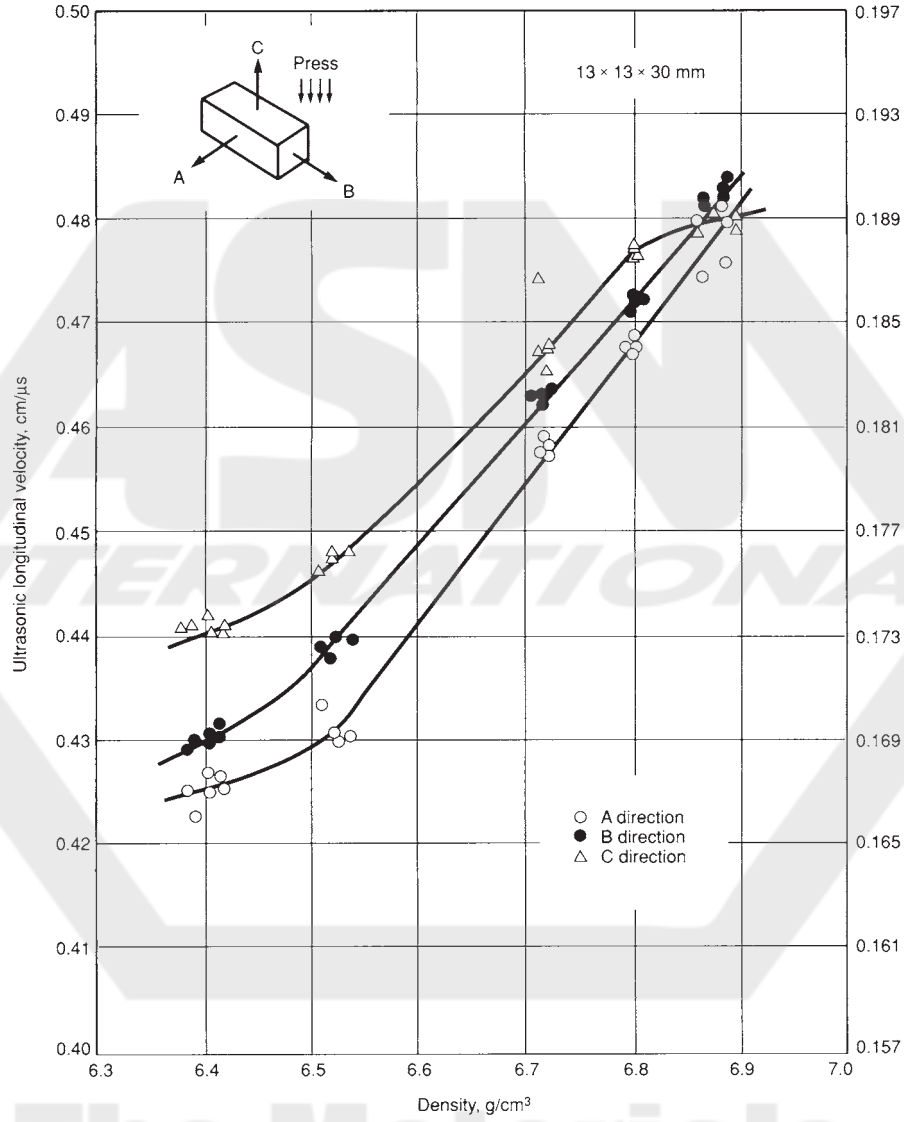


Fig. 11 Anisotropy of ultrasound velocity in sintered transverse rupture strength bars. Source: Ref 8

ACKNOWLEDGMENT

This chapter was adapted from Testing and Evaluation of Powder Metallurgy Parts in *Powder Metal Technologies and Applications*, Volume 7, *ASM Handbook*, 1998, and Nondestructive Inspection of Powder Metallurgy Parts by R.C. O'Brien and W.B. James in *Nondestructive Evaluation and Quality Control*, Volume 17, *ASM Handbook*, 1992.

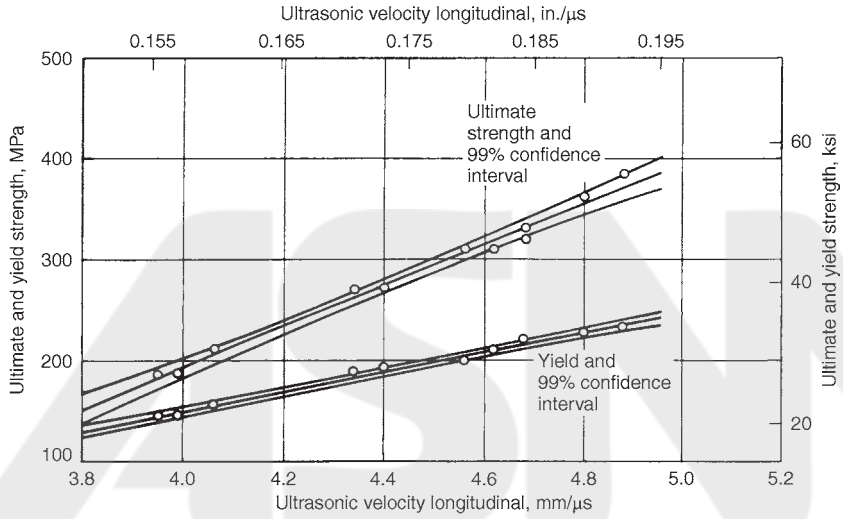


Fig. 12 Relationship of ultrasonic velocity and strength for Ancorsteel 1000B. Source: Ref 1

Table 3 Characteristics of as-sintered FN-0208 alloy specimens prepared with Ancorsteel 1000B powder, 0.75% carbon, and 0.75% Acrawax C

Sintering time, min	Final density, g/cm³	Ultrasonic velocity		Yield strength		Tensile strength	
		mm/μs	in./μs	MPa	ksi	MPa	ksi
5	6.26	3.95	0.1557	143	20.8	182	26.4
15	6.27	3.98	0.1568	142	20.6	182	26.4
30	6.27	4.06	0.1600	154	22.4	210	30.4
30	6.29	…	…	…	…	…	…
30	6.45(a)	4.42	0.1741	…	…	273	39.6
45	6.48	4.35	0.1711	183	26.6	266	38.6
5	6.76	4.56	0.1795	200	29.0	307	44.6
15	6.74	4.63	0.1821	210	30.4	307	44.6
30	6.74	4.69	0.1847	220	32.0	330	47.8
30	6.77	…	…	…	…	…	…
…	6.81(a)	4.88	0.1922	234	34.0	379	55.0
30	6.75	4.69	0.1846	218	31.6	317	46.0
45	6.89	4.83	0.1902	232	33.6	363	52.6

(a) Restruck. Source: Ref 1

REFERENCES

1. Testing and Evaluation of Powder Metallurgy Parts, *Powder Metal Technologies and Applications*, Vol 7, *ASM Handbook*, ASM International, 1998, p 710 – 718
2. R.C. O'Brien and W.B. James, Nondestructive Inspection of Powder Metallurgy Parts, *Nondestructive Evaluation and Quality Control*, Vol 17, *ASM Handbook*, ASM International, 1992, p 536–548
3. A. Lewis, "Nondestructive Inspection of Powder Metallurgy Parts Through the Use of Resistivity Measurements," Paper presented at the

Prevention and Detection of Cracks in Ferrous P/M Parts Seminar, Metal Powder Industries Federation, 1988
4. C. Rain, Uncovering Hidden Flaws, *High Technol.*, Feb 1984
5. B. Chang et al., Spatial Resolution in Industrial Tomography, *IEEE Trans. Nuclear Sci.*, NS30 (No. 2), April 1983
6. G. Schlieper, W.J. Huppmann, and A. Kozuch, Nondestructive Determination of Sectional Densities by the Gamma Densomat, *Prog. Powder Metall.*, Vol 43, 1987, p 351
7. R.H. Brockelman, Dynamic Elastic Determination of the Properties of Sintered Powder Metals, *Perspect. Powder Metall.*, Vol 5, 1970, p 201
8. E.P. Papadakis and B.W. Petersen, Ultrasonic Velocity as a Predictor of Density in Sintered Powder Metal Parts, *Mater. Eval.*, April 1979, p 76

CHAPTER 17

Inspection of Weldments and Brazed Assemblies

THE SELECTION of a method for inspecting weldments and brazed assemblies for flaws (referred to as discontinuities in welding terminology) depends on a number of variables, including the nature of the discontinuity, the accessibility of the joint, the types of materials joined, the number of joints to be inspected, the detection capabilities of the inspection method, the level of joint quality required, and economic considerations. Regardless of the method selected, established standards must be followed to obtain valid inspection results.

In general, nondestructive inspection (NDI) methods are preferred over destructive inspection methods. Sections can be trepanned from a joint to determine its integrity; however, the joint must be refilled, and there is no certainty that discontinuities would not be introduced during repair. Destructive inspection is usually impractical, because of the high cost and the inability of such methods to accurately predict the quality of those joints that were not inspected.

Weldments

Weldments made by the various welding processes may contain discontinuities that are characteristic of that process. Therefore, each process, as well as the discontinuities typical of that process, are discussed in this section.

Discontinuities in Arc Welds

Discontinuities may be divided into three broad classifications: design related, welding process related, and metallurgical. Design related discontinuities include problems with design or structural details, choice of the

wrong type of weld joint for a given application, or undesirable changes in cross section.

Discontinuities resulting from the welding process include:

- *Undercut:* A groove melted into the base metal adjacent to the toe of a weld and left unfilled by weld metal
- *Slag inclusions:* Nonmetallic solid materials entrapped in the weld metal or between the weld metal and the base metal
- *Porosity:* Cavity type discontinuities formed by gas entrapment during solidification
- *Overlap:* The protrusion of weld metal beyond the toe or root of the weld
- *Tungsten inclusions:* Particles from tungsten electrodes embedded in the weld metal that result from improper gas tungsten arc welding procedures
- *Backing piece left on:* Failure to remove material placed at the root of a weld joint to support molten weld metal
- *Shrinkage voids:* Cavity-type discontinuities normally formed by shrinkage during solidification
- *Oxide inclusions:* Particles of surface oxides that have not melted and are mixed into the weld metal
- *Incomplete fusion (also known as lack of fusion):* A condition in which fusion of the weld metal to the base metal is less than complete
- *Incomplete penetration (also known as lack of penetration):* A condition in which joint penetration is less than that specified
- *Craters:* Depressions at the termination of a weld bead or in the molten weld pool
- *Melt-through:* A condition resulting when the arc melts through the bottom of a joint welded from one side
- *Spatter:* Metal particles expelled during welding and deposited on the base metal surface
- *Arc strikes (arc burns):* Discontinuities consisting of any localized remelted metal, heat affected metal, or change in the surface profile of any part of a weld or base metal resulting from an arc
- *Underfill:* A depression on the face of the weld or root surface extending below the surface of the adjacent base metal

Metallurgical discontinuities include:

- *Cracks:* Fracture type discontinuities characterized by a sharp tip and high ratio of length and width to opening displacement
- *Fissures:* Small cracklike discontinuities with only a slight separation (opening displacement) of the fracture surfaces
- *Fisheye:* A discontinuity found on the fracture surface of a weld in steel that consists of a small pore or inclusion surrounded by a bright, round area

- *Segregation:* The nonuniform distribution or concentration of impurities or alloying elements that arises during the solidification of the weld
- *Lamellar tearing:* A type of cracking that occurs in the base metal or heat affected zone (HAZ) of restrained weld joints that is the result of inadequate ductility in the through-the-thickness direction of steel plate

The observed occurrence of discontinuities and their relative amounts depend largely on the welding process used, the inspection method applied, the type of weld made, the joint design and fit-up obtained, the material utilized, and the working and environmental conditions. The most frequent weld discontinuities found during manufacture, ranked in order of decreasing occurrence on the basis of arc welding processes, are:

Shielded metal arc welding (SMAW)	**Flux cored arc welding (FCAW)**
Slag inclusions	Slag inclusions
Porosity	Porosity
Incomplete fusion/Incomplete penetration	Incomplete fusion/Incomplete penetration
Undercut	**Gas metal arc welding (GMAW)**
Submerged arc welding (SAW)	Porosity
Slag inclusions	Incomplete fusion/Incomplete penetration
Incomplete fusion/Incomplete penetration	**Gas tungsten arc welding (GTAW)**
Porosity	Porosity

The commonly encountered inclusions, as well as cracking, the most serious of weld defects, will be discussed in this section.

Gas porosity can occur on or just below the surface of a weld. Pores are characterized by a rounded or elongated teardrop shape with or without a sharp point. Pores can be uniformly distributed throughout the weld or isolated in small groups; they can also be concentrated at the root or toe of the weld. Porosity in welds is caused by gas entrapment in the molten metal by too much moisture on the base or filler metal, or by improper cleaning of the joint during preparation for welding.

The type of porosity within a weld is usually designated by the amount and distribution of the pores. Some of the types are classified as:

- *Uniformly scattered porosity:* Characterized by pores scattered uniformly throughout the weld (Fig. 1a)
- *Cluster porosity:* Characterized by clusters of pores separated by porosity free areas (Fig. 1b)
- *Linear porosity:* Characterized by pores that are linearly distributed (Fig. 1c). Linear porosity generally occurs in the root pass and is associated with incomplete joint penetration

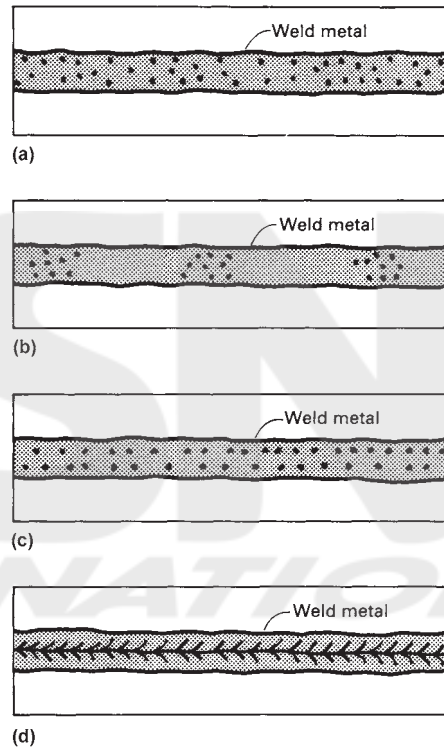


Fig. 1 Type of gas porosity commonly found in weld metal. (a) Uniformly scattered porosity. (b) Cluster porosity. (c) Linear porosity. (d) Elongated porosity. Source: Ref 1

- *Elongated porosity:* Characterized by highly elongated pores inclined to the direction of welding. Elongated porosity occurs in a herringbone pattern (Fig. 1d)
- *Wormhole porosity:* Characterized by elongated voids with a definite worm-type shape and texture (Fig. 2)

Radiography is the most widely used nondestructive method for detecting subsurface gas porosity in weldments. The radiographic image of round porosity appears as round or oval spots with smooth edges, and elongated porosity appears as oval spots with the major axis sometimes several times longer than the minor axis. The radiographic image of wormhole porosity depends largely on the orientation of the elongated cavity with respect to the incident x-ray beam. The presence of top surface or root reinforcement affects the sensitivity of inspection, and the presence of foreign material, such as loose scale, flux, or weld spatter, may interfere with the interpretation of results.

Ultrasonic inspection is capable of detecting subsurface porosity. However, it is not extensively used for this purpose except to inspect thick sections or inaccessible areas where radiographic sensitivity is limited. Surface finish and grain size affect the validity of the inspection results.

Eddy current inspection, like ultrasonic inspection, can be used for detecting subsurface porosity. Normally, eddy current inspection is confined to use on thin wall welded pipe and tubing because eddy currents are relatively insensitive to flaws that do not extend to the surface or into the near surface layer.

Magnetic particle inspection and liquid penetrant inspection are not suitable for detecting subsurface gas porosity. These methods are restricted to the detection of only those pores that are open to the surface.

Slag inclusions can occur when using welding processes that employ a slag covering for shielding purposes. With other processes, the oxide present on the metal surface before welding can also become entrapped. Slag inclusions can be found near the surface and in the root of a weld (Fig. 3a), between weld beads in multiple-pass welds (Fig. 3b), and at the side of a weld near the root (Fig. 3c).

During welding, slag may spill ahead of the arc and subsequently be covered by the weld pool because of poor joint fitup, incorrect electrode

Fig. 2 Wormhole porosity in a weld bead. Longitudinal cut. 20×. Source: Ref 1

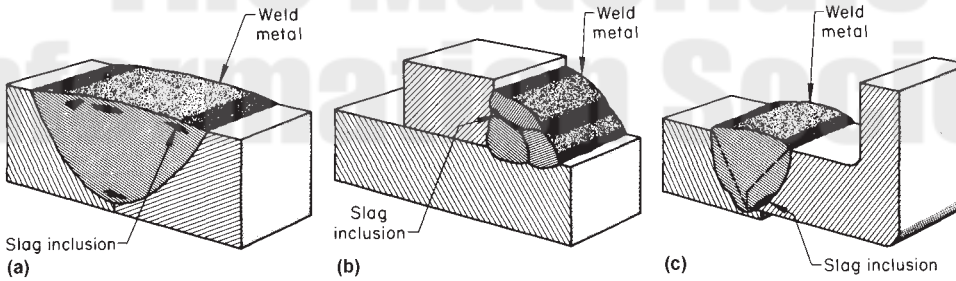


Fig. 3 Sections showing locations of slag inclusions in weld metal. (a) Near the surface and in the root of a single-pass weld. (b) Between weld beads in a multiple-pass weld. (c) At the side of a weld near the root. Source: Ref 1

manipulation, or forward arc blow. Slag trapped in this manner is generally located near the root.

Radical motions of the electrode, such as wide weaving, may also cause slag entrapment on the sides or near the top of the weld after the slag spills into a portion of the joint that has not been filled by the molten pool. Incomplete removal of the slag from the previous pass in multiple-pass welding is another common cause of entrapment. In multiple-pass welds, slag may be entrapped any number of places in the weld between passes. Slag inclusions are generally oriented along the direction of welding.

Three methods used for the detection of slag below the surface of single-pass or multiple-pass welds are magnetic particle, radiographic, and ultrasonic inspection. Depending on their size, shape, orientation, and proximity to the surface, slag inclusions can be detected by magnetic particle inspection with a dc power source, provided the material is ferromagnetic. Radiography can be used for any material, but is the most expensive of the three methods. Ultrasonic inspection can also be used for any material and is the most reliable and least expensive method. If the weld is machined to a flush contour, flaws as close as 0.8 mm ($^1/_{32}$ in.) to the surface can be detected with the straight-beam technique of ultrasonic inspection, provided the instrument has sufficient sensitivity and resolution. A 5 or 10 MHz dual element transducer is normally used in this application. If the weld cannot be machined, near surface sensitivity will be low because the initial pulse is excessively broadened by the rough, as-welded surface. Unmachined welds can be readily inspected by direct beam and reflected beam techniques, using an angle beam (shear wave) transducer.

Tungsten inclusions are particles found in the weld metal from the nonconsumable tungsten electrode used in GTAW. These inclusions are the result of:

- Exceeding the maximum current for a given electrode size or type
- Letting the tip of the electrode make contact with the molten weld pool
- Letting the filler metal come in contact with the hot tip of the electrode
- Using an excessive electrode extension
- Inadequate gas shielding or excessive wind drafts, which results in oxidation
- Using improper shielding gases such as argon-oxygen or argon-CO_2 mixtures, which are used for GMAW

Tungsten inclusions, which are not acceptable for high quality work, can only be found by internal inspection techniques, particularly radiographic testing.

Incomplete fusion and incomplete penetration result from improper electrode manipulation and the use of incorrect welding conditions. Fusion refers to the degree to which the original base metal surfaces to be

welded have been fused to the filler metal. On the other hand, penetration refers to the degree to which the base metal has been melted and resolidified to result in a deeper throat than was present in the joint before welding. In effect, a joint can be completely fused but have incomplete root penetration to obtain the throat size specified. Based on these definitions, incomplete fusion discontinuities are located on the sidewalls of a joint, and incomplete penetration discontinuities are located near the root (Fig. 4). With some joint configurations, such as butt joints, the two terms can be used interchangeably. The causes of incomplete fusion include excessive travel speed, bridging, excessive electrode size, insufficient current, poor joint preparation, overly acute joint angle, improper electrode manipulation, and excessive arc blow. Incomplete penetration may be the result of low welding current, excessive travel speed, improper electrode manipulation, or surface contaminants such as oxide, oil, or dirt that prevent full melting of the underlying metal.

Radiographic methods may be unable to detect these discontinuities in certain cases, because of the small effect they have on x-ray absorption. However, lack of sidewall fusion is readily detected by radiography.

Ultrasonically, both types of discontinuities often appear as severe, almost continuous, linear porosity because of the nature of the unbonded areas of the joint. Except in thin sheet or plate, these discontinuities may be too deep to be detected by magnetic particle inspection.

Geometric weld discontinuities are those associated with imperfect shape or unacceptable weld contour. Undercut, underfill, overlap, excessive reinforcement, fillet shape, and melt-through are included in this grouping. Geometric discontinuities are shown schematically in Fig. 5. Visual inspection is most often used to detect these flaws.

Cracks can occur in a wide variety of shapes and types and can be located in numerous positions in and around a welded joint (Fig. 6). Cracks associated with welding can be categorized according to whether they originate in the weld itself or in the base metal. Four types commonly

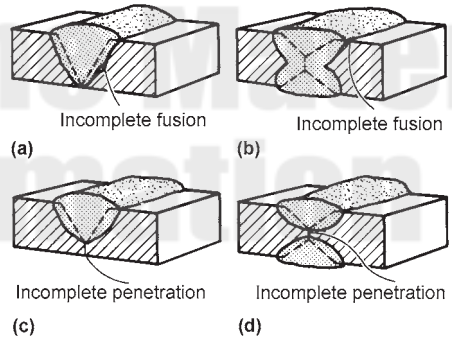


Fig. 4 Lack of fusion in (a) a single V-groove weld and (b) double V-groove weld. Lack of penetration in (c) a single V-groove and (d) a double V-groove weld. Source: Ref 1

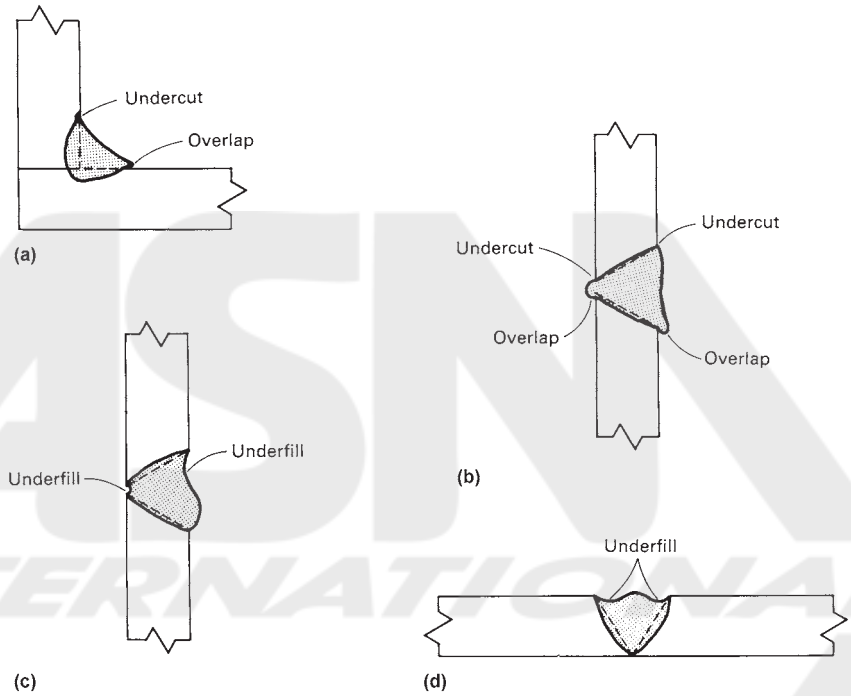


Fig. 5 Weld discontinuities affecting weld shape and contour. (a) Undercut and overlapping in a fillet weld. (b) Undercut and overlapping in a groove weld. (c) and (d) Underfill in groove welds. Source: Ref 1

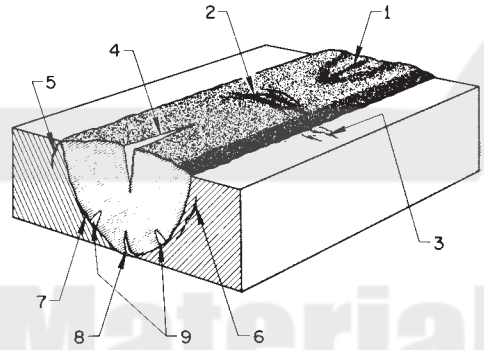


Fig. 6 Identification of cracks according to location in weld and base metal. 1, crater crack in weld metal; 2, transverse crack in weld metal; 3, transverse crack in HAZ; 4, longitudinal crack in weld metal; 5, toe crack in base metal; 6, underbead crack in base metal; 7, fusion line crack; 8, root crack in weld metal; 9, hat cracks in weld metal. Source: Ref 1

occur in the weld metal: transverse, longitudinal, crater, and hat cracks. Base metal cracks can be divided into seven categories: transverse cracks, underbead cracks, toe cracks, root cracks, lamellar tearing, delaminations, and fusion line cracks.

Weld metal cracks and base metal cracks that extend to the surface can be detected by liquid penetrant and magnetic particle inspection. Mag-

netic particle inspection can also detect subsurface cracks, depending on their size, shape, and proximity to the surface. Although the orientation of a crack with respect to the direction of the radiation beam is the dominant factor in determining the ability of radiography to detect the crack, differences in composition between the base metal and the weld metal may create shadows to hide a crack that otherwise might be visible. Ultrasonic inspection is generally effective in detecting most cracks in the weld zone.

Transverse cracks in weld metal (No. 2, Fig. 6) are formed when the predominant contraction stresses are in the direction of the weld axis. They can be hot cracks, which separate intergranularly as the result of hot shortness or localized planar shrinkage, or they can be transgranular separations produced by stresses exceeding the strength of the material. Transverse cracks lie in a plane normal to the axis of the weld and are usually open to the surface. They usually extend across the entire face of the weld and sometimes propagate into the base metal.

Transverse cracks in base metal (No. 3, Fig. 6) occur on the surface in or near the heat affected zone (HAZ). They are the result of the high residual stresses induced by thermal cycling during welding. High hardness, excessive restraint, and the presence of hydrogen promote their formation. Such cracks propagate into the weld or beyond the HAZ into the base metal as far as is needed to relieve the residual stresses.

Underbead cracks (No. 6, Fig. 6) are similar to transverse cracks in that they form in the HAZ because of high hardness, excessive restraint, and the presence of hydrogen. Their orientation follows the contour of the HAZ.

Longitudinal cracks (No. 4, Fig. 6) may exist in three forms, depending on their positions in the weld. Check cracks are open to the surface and extend only partway through the weld. Root cracks extend from the root to some point within the weld. Full centerline cracks may extend from the root to the face of the weld metal. Check cracks are caused either by high contraction stresses in the final passes applied to a weld joint or by a hot cracking mechanism.

Root cracks are the most common form of longitudinal weld metal crack because of the relatively small size of the root pass. If such cracks are not removed, they can propagate through the weld as subsequent passes are applied. This is the usual mechanism by which full centerline cracks are formed.

Centerline cracks may occur at either high or low temperatures. At low temperatures, cracking is generally the result of poor fit-up, overly rigid fit-up, or a small ratio of weld metal to base metal.

All three types of longitudinal cracks are usually oriented perpendicular to the weld face and run along the plane that bisects the welded joint. Seldom are they open at the edge of the joint face, because this requires a fillet weld with an extremely convex bead.

Crater cracks (No. 1, Fig. 6) are related to centerline cracks. As the name implies, crater cracks occur in the weld crater formed at the end of a welding pass. Generally, this type of crack is caused by failure to fill the crater before breaking the arc. When this happens, the outer edges of the crater cool rapidly, producing stresses sufficient to crack the interior of the crater. This type of crack may be oriented longitudinally or transversely or may occur as a number of intersecting cracks forming the shape of a star. Longitudinal crater cracks can propagate along the axis of the weld to form a centerline crack. In addition, such cracks may propagate upward through the weld if they are not removed before subsequent passes are applied.

Hat cracks (No. 9, Fig. 6) derive their name from the shape of the weld cross-section with which they are usually associated. This type of weld flares out near the weld face, resembling an inverted top hat. Hat cracks are the result of excessive voltage or welding speed. The cracks are located about halfway up through the weld and extend into the weld metal from the fusion line of the joint.

Toe and root cracks (No. 5 and 8, Fig. 6) can occur at the notches present at notch locations in the weld when high residual stresses are present. Both toe and root cracks propagate through the brittle HAZ before they are arrested in more ductile regions of the base metal. Characteristically, they are oriented almost perpendicular to the base metal surface and run parallel to the weld axis.

Lamellar tearing is the phenomenon that occurs in T-joints that are fillet welded on both sides. This condition, which occurs in the base metal or HAZ of restrained weld joints, is characterized by a step-like crack parallel to the rolling plane. The crack originates internally because of tensile strains produced by the contraction of the weld metal and the surrounding HAZ during cooling. A typical condition is shown in Fig. 7.

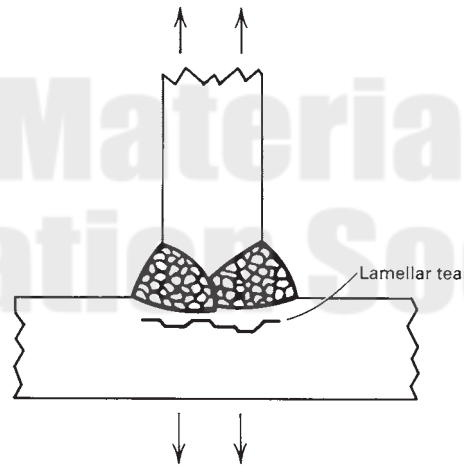


Fig. 7 Lamellar tear caused by thermal contraction strain. Source: Ref 1

Methods of Nondestructive Inspection

The nondestructive inspection of weldments has two functions:

- Quality control, which is the monitoring of welder and equipment performance and of the quality of the consumables and the base materials used
- Acceptance or rejection of a weld on the basis of its fitness for purpose under the service conditions imposed on the structure

The appropriate method of inspection is different for each function. If evaluating for discontinuities is a viable option they must be detected, identified, located exactly, sized, and their orientation established.

Weld discontinuities constitute the center of activity with the inspection of welded constructions. The most widely used inspection techniques used in the welding industry are visual, liquid penetrant, magnetic particle, radiographic, ultrasonic, acoustic emission, eddy current, and electric current perturbation methods. Each of these techniques has specific advantages and limitations. Existing codes and standards that provide guidelines for these various techniques are based on the capabilities and/or limitations of these nondestructive methods.

Selection of Technique. A number of factors influence selection of the appropriate nondestructive test technique for inspecting a welded structure, including discontinuity characteristics, fracture mechanics requirements, part size, portability of equipment, and other application constraints. These categories, although perhaps unique to a specific inspection problem, may not clearly point the way to the most appropriate technique. It is generally necessary to exercise engineering judgment in ranking the importance of these criteria and thus determining the optimum inspection technique.

Characteristics of the Discontinuity. Because nondestructive techniques are based on physical phenomena, it is useful to describe the properties of the discontinuity of interest, such as composition and electrical, magnetic, mechanical, and thermal properties. Most significant are those properties that are most different from those of the weld or base metal. It is also necessary to identify a means of discriminating between discontinuities with similar properties.

Fracture mechanics requirements, solely from a discontinuity viewpoint, typically include detection, identification, location, sizing, and orientation. In addition, complicated configurations may necessitate a nondestructive assessment of the state of the stress of the region containing the discontinuity. In the selection process, it is important to establish these requirements correctly. This may involve consultation with stress analysts, materials engineers, and statisticians. Often, the criteria may strongly suggest a particular technique. Under ideal conditions, such as in a labora-

tory, the application of such a technique might be routine. In the field, however, other factors may force a different choice of technique.

Constraints tend to be unique to a given application and may be completely different even when the welding process and metals are the same. Some of these constraints include:

- Access to the region under inspection
- Geometry of the structure (flat, curved, thick, thin)
- Condition of the surface (smooth, irregular)
- Mode of inspection (pre-service, in-service, continuous, periodic, spot)
- Environment (hostile, underwater)
- Time available for inspection (high speed, time intensive)
- Reliability
- Application of multiple techniques
- Cost

Failure to consider adequately the constraints imposed by a specific application can render the most sophisticated equipment and theory useless. Moreover, for the simple or less important cases of failure, it may be unnecessary. Once criteria have been established, an optimum inspection technique can be selected, or designed and constructed.

Visual Inspection

For many noncritical welds, integrity is verified principally by visual inspection. Even when other nondestructive methods are used, visual inspection still constitutes an important part of practical quality control. Widely used to detect discontinuities, visual inspection is simple, quick, and relatively inexpensive. The only aids that might be used to determine the conformity of a weld are a low power magnifier, a borescope, a dental mirror, or a gage. Visual inspection can and should be done before, during, and after welding. Although visual inspection is the simplest inspection method to use, a definite procedure should be established to ensure that it is carried out accurately and uniformly.

Visual inspection is useful for checking the following:

- Dimensional accuracy of weldments
- Conformity of welds to size and contour requirements
- Acceptability of weld appearance with regard to surface roughness, weld spatter, and cleanness
- Presence of surface flaws such as unfilled craters, pockmarks, undercuts, overlaps, and cracks

Although visual inspection is an invaluable method, it is unreliable for detecting subsurface flaws. Therefore, judgment of weld quality must be based on information in addition to that afforded by surface indications.

Additional information can be gained by observations before and during welding. For example, if the plate is free of laminations and properly cleaned and if the welding procedure is followed carefully, the completed weld can be judged on the basis of visual inspection. Additional information can also be gained by using other NDI methods that detect subsurface and surface flaws.

Dimensional Accuracy and Conformity. All weldments are fabricated to meet certain specified dimensions. The fabricator must be aware of the amount of shrinkage that can be expected at each welded joint, the effect of welding sequence on warpage or distortion, and the effect of subsequent heat treatment used to provide dimensional stability of the weldment in service.

Weldments that require rigid control of final dimensions usually must be machined after welding.

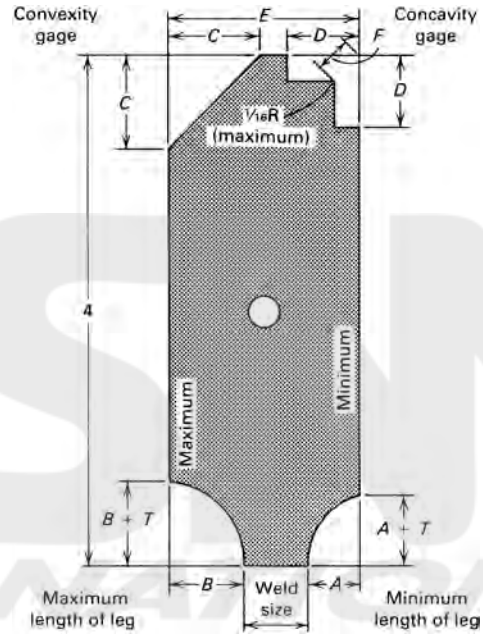
Dimensional tolerances for as-welded components depend on the thickness of the material, the alloy being welded, the overall size of the product, and the particular welding process used.

The dimensional accuracy of weldments is determined by conventional measuring methods, such as rules, scales, calipers, micrometers, and gages. The conformity of welds with regard to size and contour can be determined by a weld gage. The weld gage shown in Fig. 8 is used when visually inspecting fillet welds at 90° intersections. The size of the fillet weld, which is defined by the length of the leg, is stamped on the gage. The weld gage determines whether or not the size of the fillet weld is within allowable limits and whether there is excessive concavity or convexity. This gage is designed for use on joints between surfaces that are perpendicular. Special weld gages are used when the surfaces are at angles other than 90°. For groove welds, the width of the finished welds must be in accordance with the required groove angle, root face, and root opening. The height of reinforcement of the face and root must be consistent with specified requirements and can be measured by a weld gage.

Appearance Standards. The acceptance of welds with regard to appearance implies the use of a visual standard, such as a sample weldment or a workmanship standard. Requirements as to surface appearance differ widely, depending on the application. For example, when aesthetics are important, a smooth weld that is uniform in size and contour may be required.

The inspection of multiple-pass welds is often based on a workmanship standard. Figure 9 indicates how such standards are prepared for use in the visual inspection of groove and fillet welds. The workmanship standard is a section of a joint similar to the one in manufacture, except that portions of each weld pass are shown. Each pass of the production weld is compared with corresponding passes of the workmanship standard.

Discontinuities. Before a weld is visually inspected for discontinuities, such as unfilled craters, surface holes, undercuts, overlaps, surface cracks,



A	Minimum allowable length of leg
B	Maximum allowable length of leg
C	1.414 times maximum allowable throat size (specifies maximum allowable convexity)
D	Maximum allowable length of leg when maximum allowable concavity is present
E	*A* plus *B* plus nominal weld size (or nominal length of leg)
F	Minimum allowable throat size (specifies maximum allowable concavity)
T	Additional tolerance for clearance of gage when placed in the fillet

Fig. 8 Gage for visual inspection of a fillet weld at a 90° intersection. Similar gages can be made for other angles. Dimensions given in inches. Source: Ref 1

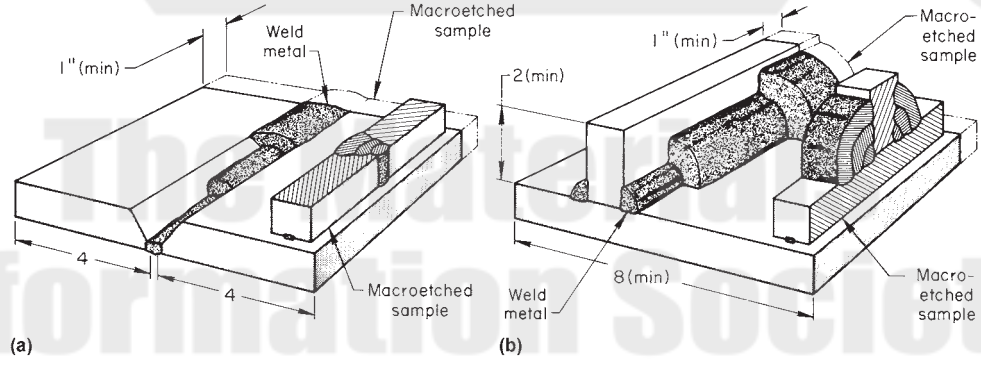


Fig. 9 Workmanship standard for visual comparison during inspection of (a) single V-groove welds and (b) fillet welds. Dimensions given in inches. Source: Ref 1

and incomplete joint penetration, the surface of the weld should be cleaned of oxides and slag. Cleaning must be done carefully. For example, a chipping hammer used to remove slag could leave hammer marks that can hide fine cracks. Shot blasting can peen the surface of relatively soft metals and hide flaws. A stiff wire brush and sandblasting have been found to be satisfactory for cleaning surfaces of slag and oxides without marring.

Magnetic Particle Inspection

Magnetic particle inspection is particularly suitable for the detection of surface flaws in highly ferromagnetic metals. Under favorable conditions, those discontinuities that lie immediately under the surface are also detectable. Nonferromagnetic and weakly ferromagnetic metals, which cannot be strongly magnetized, cannot be inspected by this method. With suitable ferromagnetic metals, magnetic particle inspection is highly sensitive and produces readily discernible indications at flaws in the surface of the material being inspected.

The types of weld discontinuities normally detected by magnetic particle inspection include cracks, incomplete penetration, incomplete fusion, and porosity open to the surface. Linear porosity, slag inclusions, and gas pockets can be detected if large or extensive or if smaller and near the surface. The recognition of patterns that indicate deep-lying flaws requires more experience than that required to detect surface flaws.

Nonrelevant indications that have no bearing on the quality of the weldment may be produced. These indications are magnetic particle patterns held by conditions caused by leakage fields. Some of these conditions are:

- Particles held mechanically or by gravity in surface irregularities
- Adherent scale or slag
- Indications at a sharp change in material direction, such as sharp fillets and threads
- Grain boundaries. Large grain sizes in the weld metal or the base metal may produce indications
- Boundary zones in welds, such as indications produced at the junction of the weld metal and the base metal. This condition occurs in fillet welds at T-joints, or in double V-groove joints, where 100% penetration is not required
- Flow lines in forgings and formed parts
- Brazed joints. Two parts made of a ferromagnetic material joined by a nonferromagnetic material will produce an indication
- Different degrees of hardness in a material, which will usually have different permeabilities that may create a leakage field, forming indications

Operational Requirements. The magnetic particle inspection of weldments requires that the weld bead be free of scale, slag, and moisture. For maximum sensitivity, the weld bead should be machined flush with the surface; however, wire brushing, sandblasting, or grit blasting usually produces a satisfactory bead surface. If the weld bead is rough, grinding will remove the high spots.

Weldments are often inspected using the dry particle method. A powder or paste of a color that gives the best possible contrast to the surface being inspected should be used. The type of magnetizing current used depends on whether there are surface or subsurface discontinuities. Alternating current is satisfactory for surface cracks, but if the deepest possible penetration is essential, direct current, direct current with surge, or half-wave rectified alternating current is used.

The voltage should be as low as practical to reduce the possibility of damage to the surface of the part from overheating or arcing at contacts. Another advantage of low voltage is freedom from arc flashes if a prod slips or is withdrawn before the current is turned off.

The field strength and flux density used must be determined for each type of weldment. An overly strong field will cause the magnetic particles to adhere too tightly to the surface and hinder their mobility, preventing them from moving to the sites of the flaws. Low field strengths result in nondiscernible patterns and failure to detect indications.

Inspections can be made using the continuous-field and residual-field methods. In the continuous-field method, magnetic particles are placed on the weldment while the current is flowing. In the residual-field method, the particles are placed on the weldment after the magnetizing current is turned off. Residual magnetic fields are weaker than continuous fields. Consequently, inspections using the residual-field method are less sensitive.

The need for the demagnetization of weldments after magnetic particle inspection must be given serious consideration. Where subsequent welding or machining operations are required, it is good practice to demagnetize. Residual magnetism may also hinder cleaning operations and interfere with the performance of instruments used near the weldment.

Liquid Penetrant Inspection

Liquid penetrant inspection is capable of detecting discontinuities open to the surface in weldments made of either ferromagnetic or nonferromagnetic alloys, even when the flaws are generally not visible to the unaided eye. For the correct usage of liquid penetrant inspection, it is essential that the surface of the part be thoroughly clean, leaving the openings free to receive the penetrant. Operating temperatures of 20 to 30 °C (70 to 90 °F) produce optimum results. If the part is cold, the penetrant may become chilled and thickened so that it cannot enter very fine openings. If the part

or the penetrant is too hot, the volatile components of the penetrant may evaporate, reducing the sensitivity.

After the penetrating period, the excess penetrant remaining on the surface is removed. An absorbent, light colored developer is then applied to the surface. This developer acts as the blotter, drawing out a portion of the penetrant that had previously seeped into the surface openings. As the penetrant is drawn out, it diffuses into the developer, forming indications that are wider than the surface openings. The inspector looks for these colored or fluorescent indications against the background of the developer.

Radiographic Inspection

Surface discontinuities that are detectable by radiography include undercuts, longitudinal grooves, concavity at the weld root, incomplete filling of grooves, excessive penetration, offset or mismatch, burn-through, irregularities at electrode change points, grinding marks, and electrode spatter. Surface irregularities may cause density variations on a radiograph. When possible, they should be removed before a weld is radiographed. When impossible to remove, they must be considered during interpretation.

Undercuts result in a radiographic image of a dark line of varying width and density. The darkness or density of the line indicates the depth of the undercut.

Longitudinal grooves in the surface of weld metal produce dark lines on a radiograph that are roughly parallel to the weld seam but are seldom straight. These dark lines have diffused edges and should not be mistaken for slag lines, which are narrow and more sharply defined.

Concavity at the weld root occurs only in joints that are welded from one side, such as pipe joints. It appears on the radiograph as a darker region than the base metal.

If weld reinforcement is too high, the radiograph shows a lighter line down the weld seam. There is a sharp change in image density where the reinforcement meets the base metal. Weld reinforcements not ground completely smooth show irregular densities, often with sharp borders.

When excess metal is deposited on a final pass, it may overlap the base metal, causing incomplete fusion at the edge of the reinforcement. Although there is a sharp change in image density between reinforcement and base metal, the edge of the reinforcement image is usually irregular.

Irregularities at electrode change points may be either darker or lighter than the adjacent areas.

Grinding marks appear as darker areas or lines in relation to the adjacent areas in the radiograph.

Electrode spatter will appear as globular and lighter on the radiograph and should be removed before radiographic inspection.

As material thickness increases, radiography becomes less sensitive as an inspection method. Thus, for thick material, other NDI methods are used before, during, and after welding on both the base metal and weld metal.

Subsurface discontinuities detectable by radiography include gas porosity, slag inclusions, cracks, incomplete penetration, incomplete fusion, and tungsten inclusions. On a radiograph, a pore appears as a round or oval dark spot with or without a rather sharp tail. The spots caused by porosity are often of varying size and distribution. A wormhole appears as a dark rectangle if its long axis is perpendicular to the radiation beam, and it appears as two concentric circles, one darker than the other, if the long axis is parallel to the beam. Linear porosity is recorded on radiographs as a series of round dark spots along a line parallel to the direction of welding.

Slag inclusions appear along the edge of a weld as irregular or continuous dark lines on the radiograph. Voids are sometimes present between weld beads because of irregular deposition of metal during multiple-pass welding. These voids have a radiographic appearance that resembles slag lines.

The radiographic image of a crack is a dark narrow line that is generally irregular. If the plane of the crack is in line with the radiation beam, its image is a fairly distinct line. If the plane is not exactly in line with the radiation beam, a faint dark linear shadow results. In this case, additional radiographs should be taken at other angles.

Incomplete penetration shows on a radiograph as a very narrow dark line near the center of the weld. The narrowness can be caused by drawing together of the plates being welded, and the incomplete penetration may be very severe. Slag inclusions and gas holes are sometimes found in connection with incomplete penetration and cause the line to appear broad and irregular.

The radiographic image of incomplete fusion shows a very thin, straight dark line parallel to and on one side of the weld image. Where there is doubt, additional radiographs should be made with the radiation beam parallel to the bevel face. This will increase the possibility of the incomplete fusion appearing on the radiograph.

Tungsten inclusions appear either as single light spots or as clusters of small light spots. The spots are usually irregular in shape, but sometimes a rectangular light spot will appear.

Real-time radiography, which involves the display of radiographic images on television monitors through the use of an image converter and a television camera, is a rapidly developing method for weld inspection. One of the main advantages of real-time radiography for weld inspection is the cost savings that results from reducing the use of x-ray films. However, the possibility of expanding such an inspection system to include automatic defect evaluation by the image processing system can yield sig-

nificantly greater advantages. Automatic defect evaluation systems will result in objective and reproducible x-ray inspection, independent of human factors.

Until now, the human brain has been much faster in analyzing and classifying the large range of flaw types found in welded joints. Computer programs for the efficient automated evaluation of weld radiographs are currently being developed and refined.

Ultrasonic Inspection

In ultrasonic inspection, a beam of ultrasonic energy is directed into a specimen, and either the energy transmitted through the specimen is measured or the energy reflected from interfaces is indicated. Normally, only the front (entry) and back surfaces plus discontinuities within the metal produce detectable reflections, but in rare cases, the HAZs or the weld itself may act as reflecting interfaces.

Scanning Techniques. Figure 10 shows how a shear wave from an angle beam transducer progresses through a flat test piece, by reflecting from the surfaces at points called nodes. The linear distance between two successive nodes on the same surface is called the skip distance and is important in defining the path over which the transducer should be moved for reliable and efficient scanning of a weld. The skip distance can be easily measured by using a separate receiving transducer to detect the nodes, or by using an angle beam test block, or it can be calculated. Once the skip distance is known, the region over which the transducer should be moved to scan the weld can be determined. This region should extend the entire length of the weld at a distance from the weld line of approximately ½ to 1 skip distance, as shown in Fig. 11. A zigzag scanning path is used, either with sharp changes in direction (Fig. 11) or with squared changes (Fig. 12).

To detect longitudinal discontinuities in full penetration butt and corner welds that are not ground flush, the transducer is oscillated to the left and

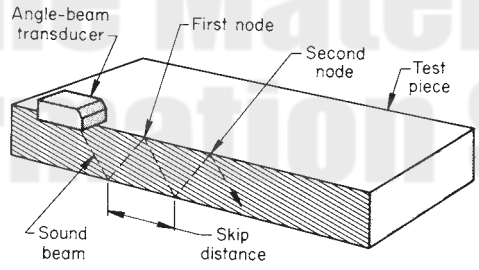


Fig. 10 Sound beam path in a flat test piece being ultrasonically inspected with a shear wave from an angle beam transducer, showing the skip distance between the nodes where the beam reflects from the surfaces. Source: Ref 1

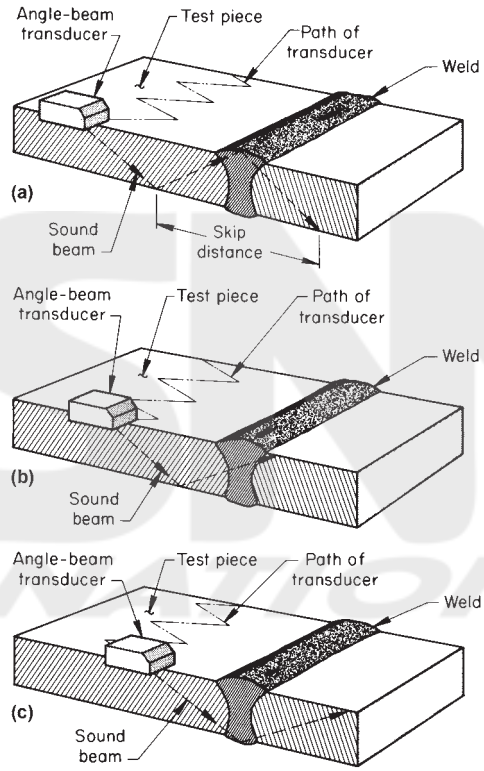


Fig. 11 Three positions of the contact type of transducer along the zigzag scanning path used during ultrasonic inspection of welded joints. The movement of the sound beam path across the weld is shown on a section taken along the centerline of the transducer as it is moved (a) from the far left position in the scanning path, (b) through an intermediate position, (c), to the far right position. Source: Ref 1

right in a radial motion, with an included angle of approximately 30°, while scanning perpendicularly toward the weld, as shown in Fig. 12(a). The longitudinal movement necessary to advance the transducer parallel to the weld should not exceed 75% of the active width of the transducer per transverse scan. The weld should be scanned from both sides on one surface or from one side on both surfaces to ensure that nonvertically-oriented flat discontinuities are detected. This type of discontinuity can be distinguished from vertically oriented flat discontinuities because the signal amplitudes from the two sides are different.

To detect transverse discontinuities in welds that are not ground flush, the transducer is placed on the base metal surface at the edge of the weld. The sound beam is directed by angling the transducer approximately 15% toward the weld from the longitudinal weld axis, as shown in Fig. 12(a). Scanning is performed by moving the transducer along the edge of the weld either in one direction along both sides of the weld or in opposite directions along one side of the weld.

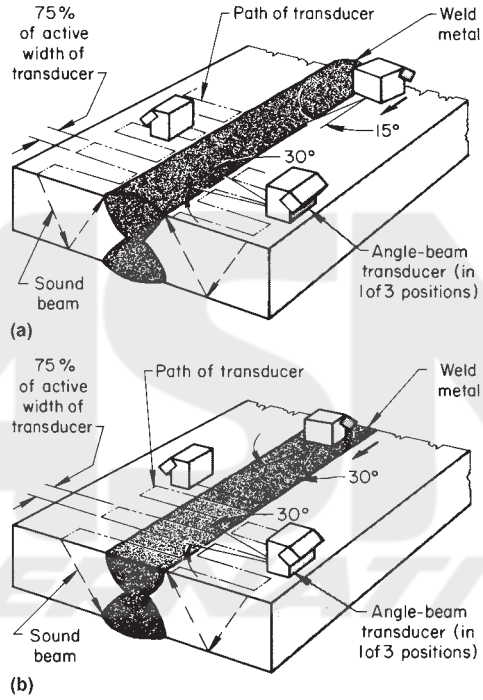


Fig. 12 Ultrasonic scanning procedures to detect longitudinal and transverse discontinuities in welds that (a) are not ground flush and (b) are ground flush. Source: Ref 1

To detect longitudinal discontinuities in welds that are ground flush, the transducer is oscillated to the left and right in a radial motion, with an included angle of approximately 30°, while scanning across the weld as shown in Fig. 12(b). The longitudinal movement necessary to advance the transducer parallel to the weld must not exceed 75% of the active width of the transducer per transverse scan. When possible, the weld is scanned from one surface on two sides of the weld.

When this is not possible, the weld can be scanned from one side on two surfaces or from one side on one surface using at least one full skip distance.

To detect transverse discontinuities in welds that are ground flush, the transducer is oscillated to the left and right in a radial motion, with an included angle of approximately 30°, as shown in Fig. 12(b), while scanning along the top of the weld from two opposing directions. If the width of the weld exceeds the width of the transducer, parallel scans should be performed, with each succeeding scan overlapping the previous one by a minimum of 25% of the active width of the transducer.

The entire volume of full penetration welds in corner joints should be scanned with shear waves by directing the sound beam toward, or across and along, the axis of the weld, as shown in Fig. 13. If longitudinal wave

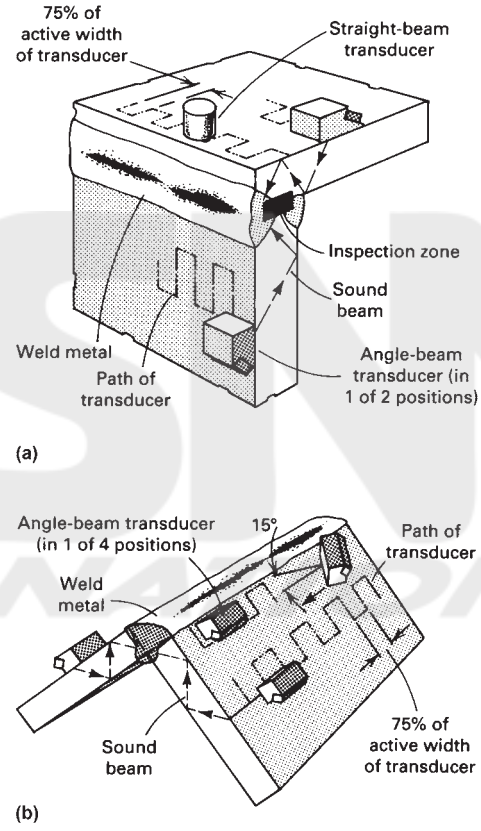


Fig. 13 Ultrasonic scanning procedure for full penetration groove weld (a) and double fillet welds (b) in corner joints. Source: Ref 1

testing is utilized, the weld is scanned by moving the transducer over the weld with overlapping paths. Each succeeding scan should overlap the previous scan by at least 25% of the active width of the transducer.

For the detection of discontinuities in the root area in T-joints (such as lack of root fusion), the width of the inspection zone should be limited to the thickness of the attachment member. The width of the inspection zone is located using ultrasonics or mechanical means and marked on the test surface. Shear wave scanning for discontinuities in the base metal of any T-joint configuration should be performed whenever the surface opposite the attachment member is accessible. This scanning procedure can also be applied to partial penetration welds in T-joints. Coverage in each direction begins from the nearest section of the joint to beyond the centerline of the weld. The angle beam transducer is directed at the particular area of interest and oscillated to the left and right in a radial motion, with an included angle of approximately 30°, while scanning perpendicularly toward the inspection zone. The inspection zone depth should be limited to the through member plate thickness minus 6 mm (¼ in.). The movement nec-

essary to advance the transducer parallel to the inspection zone should not exceed 75% of the active width of the transducer per perpendicular scan.

Discontinuity Signals. Cracks and incomplete fusion discontinuities present essentially flat reflectors to the ultrasonic beam. If the beam is perpendicular to the plane of the discontinuity, the amplitude of the signal is high; but if the beam strikes the discontinuity at an angle, most of the ultrasonic energy is reflected away from the transducer, and the reflected signal has a small amplitude that will vary with the angle. Because both cracks and sidewall incomplete fusion discontinuities produce similar reflected signals, they cannot be distinguished from one another by the signal amplitude or signal shape on the viewing screen when scanning is done from only one side. Therefore, the weld should be inspected from two sides, in the manner shown in Fig. 14. If the discontinuity is vertically oriented, such as a centerline crack would be, the reflected signals received during a scan of each side should have approximately the same amplitude. If the discontinuity is in an inclined position, such as a sidewall incomplete fusion discontinuity would be in many joint designs, there will be an appreciable difference between the signal amplitudes.

A slag inclusion in a butt weld may produce a reflected signal with the same amplitude as that received from a crack or incomplete fusion discontinuity. However, scattered ultrasonic energy produces a relatively wide and high signal; as the transducer is manipulated around the slag inclusion, the signal height does not decrease significantly, but the edges of the signal vary. The same shape of the reflected signal should be displayed when the weld is scanned from the opposite side of the weldment. The signals that are reflected from porosity (gas pockets) are usually small and narrow. The signal amplitude will vary if the transducer is manipulated around the gas pocket or if the gas pocket is scanned from the opposite side of the weld.

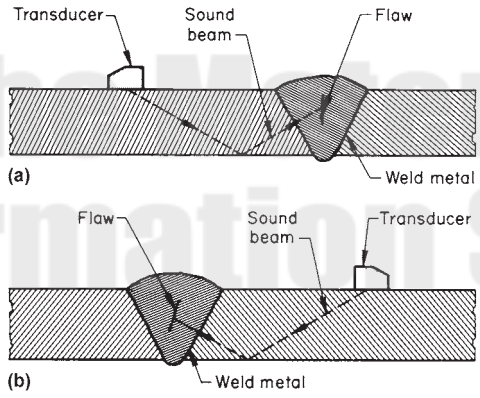


Fig. 14 Transducer scanning positions for distinguishing between weld metal flaws that are (a) vertically oriented and (b) in an inclined position. Source: Ref 1

Cluster porosity (groups of gas pockets) usually produces displays with a number of small signals. Depending on the number of gas pockets and their orientation to the ultrasonic beam, the displayed signals will be stationary or will be connected with one another.

Lack of fusion, weld root cracks, and incomplete penetration give essentially the same type of signal on an oscilloscope screen; the reflected signals are narrow and appear at the same location. The best way to differentiate among these flaws is to determine the extent of the flaw in the transverse direction.

Weld undercutting is distinguishable from sidewall incomplete fusion. The signals reflected from undercutting are approximately equal in amplitude when scanned from both sides. The signals produced by a sidewall incomplete fusion discontinuity vary considerably in amplitude when scanned from both sides.

In many cases, a weld is made when two misaligned parts must be joined; this condition is termed *weld mismatch*. The inspector must not confuse a signal reflected from a root crack with one reflected from the misaligned edge. A narrow signal is usually produced when the ultrasonic beam strikes the misaligned edge. In most cases, no reflected signal will be received if the misaligned edge is scanned from the opposite side.

Ultrasonic Inspection of Spot Welds in Thin Gage Steel. With the development of high frequency transducers (12 to 20 MHz), the pulse echo ultrasonic inspection of spot welds in very thin gage sheet metal (0.58 mm, or 0.023 in.) is now possible. The ultrasonic test for spot weld nugget integrity relies on an ultrasonic wave to measure the size of the nugget. The size is in three dimensions, including thickness as well as length and breadth (or diameter for a circular spot). The successful measurement of nugget size places several requirements on the ultrasonic wave path, wave velocity, and wave attenuation.

Wave Path. The first requirement is that the ultrasonic wave be in the form of a beam directed perpendicular to the faces of the metal sheets and through the center of the nugget (Fig. 15). Two diameters of nuggets are shown: larger than the beam and smaller than the beam.

In general, an ultrasonic wave will be reflected when it impinges on an interface where the density and/or the ultrasonic velocity change. Examples are water-to-metal and metal-to-air. In Fig. 15, reflections will occur at the outer surfaces of the two sheets and at the interface (air) between the two sheets if the nugget is small, as in Fig. 15(c). The nugget-to-parent metal boundary will not produce perceptible reflections, refraction, or scattering, because the changes in density and velocity are a tenth of a percent or less, while the air-to-steel difference is more than 99.9%. Typical oscilloscope displays showing the pulse echo patterns for these two nugget-to-beam diameter ratios are shown in Fig. 15(b) and 15(d). The difference in the echo patterns permits the distinction to be made between adequate and undersize welds.

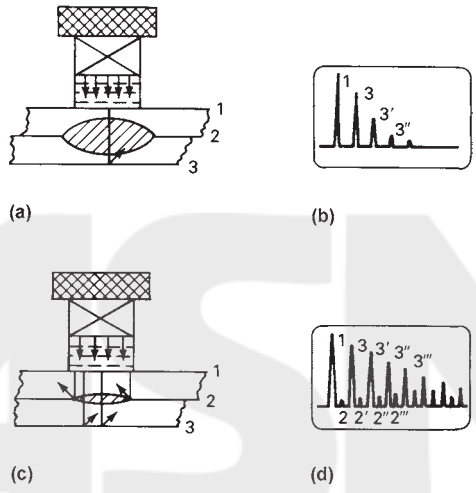


Fig. 15 Schematic illustrating setup for the pulse echo ultrasonic inspection of resistance welded spot welds. (a) Wave paths in satisfactory weld. (b) Resulting echoes. (c) Wave paths in an unsatisfactory weld. (d) Resulting echoes. Source: Ref 1

Velocity/Thickness Gaging. The beam path shown in Fig. 15(a) illustrates the situation in which the ultrasonic beam should indicate an acceptable nugget. The beam will be reflected only at the outer surfaces (1 and 3) of the pair of sheets as joined. To make this reflection sequence visible, the ultrasonic beam must consist of a short pulse that can reverberate back and forth between the outer faces and produce separate echoes when viewed on an oscilloscope. The picture observed is illustrated in Fig. 15(b). The pulse must be short enough to resolve the double thickness of the two joined sheets.

Similarly, the beam path shown in Fig. 15(c) illustrates the situation in which the ultrasonic beam should indicate an undersize nugget. The beam will be reflected in the single thickness of the upper sheet around the perimeter of the nugget. Therefore, on the oscilloscope, echoes will appear between the principal echoes arising from the portion of the beam traversing the nugget (Fig. 15d). In terms of thickness gaging, the ultrasonic pulse in the beam must be short enough to resolve the thickness of one layer of sheet metal.

Attenuation. The thickness of the nugget can only be measured indirectly because the thickness gaging function can measure only the thickness between outer faces in the nugget area. The nugget itself is measured by the effect of its grain structure on the attenuation of the ultrasonic wave in the beam. As the wave reflects back and forth between the outer faces of the welded sheets, its amplitude is *attenuated* or dies out. The *attenuation* or rate of decay of the ultrasonic wave depends on the microstructure of the metal in the beam. In the spot welds under consideration, the attenuation is caused principally by grain scattering. The grains scatter the ultra-

sonic energy out of the coherent beam, causing the echoes to die away. In most metals, coarse grains scatter more strongly than fine grains.

Because a nugget is a melted and subsequently refrozen cast microstructure with coarser grains than the adjacent cold rolled parent metal, the nugget will scatter more strongly than the remaining parent metal. It follows that a nugget will produce higher attenuation than the parent metal and that a thick nugget will result in higher attenuation than will a thin nugget. Therefore, a thin nugget can be distinguished from a thick nugget by the rate of decay of the echoes in the case in which the diameters of both nuggets are equal. Typical echoes from a thick nugget area and from a thin nugget are shown in Fig. 16. A trained observer can differentiate between the two welds on the basis of the decay patterns.

Given this observation, it is obvious that the pulse echo ultrasonic method at normal incidence could perform the required measurements on spot welds in metals with coarse grain nuggets and fine grain parent sheet metal.

Leak Testing

Welded structures are leak tested to measure the integrity of the structure for containing gases, fluids, semisolids, and solids and for maintaining pressures and vacuums. The more common leak testing methods used, in order of increasing sensitivity, are:

- Odor from tracer gas
- Pressure change
- Pressurized liquid (generally water) and visual observation

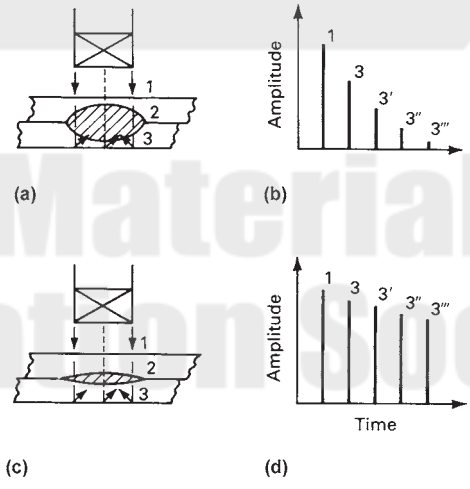


Fig. 16 Ultrasonic thickness measurements of resistance spot weld nuggets. (a) Satisfactory weld. (b) Resulting attenuation of the ultrasonic wave. (c) Unsatisfactory weld. (d) Resulting wave attenuation. Source: Ref 1

- Pressurized gas using a leak detection solution
- Tracer gas using thermal leak detectors
- Helium using a mass spectrometer during pressure and vacuum tests

Other methods less frequently used are acoustical detection of gas flow through a leak and use of radioactive tracer gas.

Weld flaws that contribute to leakage of a structure are porosity, incomplete fusion or incomplete penetration, and cracks. Cracks are of particular concern because they may propagate when the structure is proof tested or otherwise tested for structural integrity. Therefore, it is preferred that leak testing be done after completion of the structural tests.

Selection of a leak testing method depends on the environment in which the structure is used and the potential danger and economic impact involved in the event of a service failure. The acceptance criteria should include a numerical expression of the allowable leak rate; the frequently used expression "shall be free from leaks" is meaningless.

When conducting pressure tests with compressible gases (e.g., air), extreme caution is necessary. If a pressure vessel that is pressurized with a compressible gas fails during a leak or proof test, an explosion can occur. In general, most pressurized proof tests are conducted with incompressible fluids, such as water. In this case, if failure under pressure occurs, a leak rather than explosion will occur.

Eddy Current and Electric Current Perturbation Inspection

Eddy current inspection is based on the principles of electromagnetic induction and is used to identify or to differentiate between a wide variety of physical, structural, and metallurgical conditions in electrically conductive ferromagnetic and nonferromagnetic metals. Normally, eddy current inspection is used only on thin wall welded pipe and tubing for the detection of longitudinal weld discontinuities, such as open welds, weld cracks, and porosity.

The electric current perturbation method consists of establishing an electric current flow in the part to be inspected (usually by means of an induction coil) and detecting the magnetic field associated with perturbations in the current flow around defects by using a separate magnetic field sensor. This technique is applicable to the detection of both very small surface cracks as well as subsurface cracks in both high and low conductivity, nonferromagnetic materials such as titanium and aluminum alloys.

Brazed Assemblies

Brazing is defined by the American Welding Society as a group of welding processes that produce coalescence of materials by heating them to a suitable temperature and by using a filler metal having a liquidus above

450 °C (840 °F) and below the solidus of the base metal. The filler metal is distributed between the closely fitted faying surfaces of the joint by capillary action.

The temperature limitation of 450 °C (840 °F) differentiates brazing from soft soldering, which involves the use of filler metals having a liquidus below 450 °C (840 °F). To clarify the difference between brazing and conventional welding, it should be pointed out that in brazing the base materials being joined are never melted, while in most welding processes the base metals are melted (exceptions are those welding processes that utilize pressure in conjunction with heat).

There are six brazing processes included under the group heading of brazing. These processes are torch brazing, furnace brazing, induction brazing, dip brazing, resistance brazing, and infrared brazing.

Brazing can also be classified according to the major constituents of the more common types of filler metals used:

- Aluminum brazing
- Silver brazing
- Copper brazing
- Nickel brazing
- Precious metal brazing

The five essential properties for brazing filler metal are:

- Ability to wet and make a strong, sound bond on the base metal
- Suitable melting temperature and flow properties to permit distribution by capillary attraction in properly prepared joints
- A composition of sufficient homogeneity and stability to minimize separation by liquation under the brazing conditions to be encountered. Excessively volatile constituents in filler metals may be objectionable
- Capability of producing a brazed joint that will meet service requirements, such as required strength and corrosion resistance
- Depending on the requirements, ability to produce or avoid interactions between the base metal and filler metal

Flaws Commonly Found in Brazed Joints

The usual types of flaws exhibited by brazed joints are:

- Lack of fill
- Flux entrapment
- Noncontinuous fillet
- Base metal erosion

Lack of Fill. Voids resulting from lack of fill can be the result of improper cleaning of the faying surfaces, improper clearances, insufficient brazing temperatures, or insufficient brazing filler metal (Fig. 17).

Flux entrapment normally occurs during torch brazing, induction brazing, or furnace brazing, when reducing atmospheres are not employed. As the term implies, flux becomes trapped within the joint by the braze metal and prevents coverage. Figure 18 is a radiograph of a torch brazed joint in which flux entrapment was a serious problem.

Noncontinuous Fillet. A brazed joint in which a large void in the fillet is evident is shown in Fig. 19. Such a void is discernible by visual examination and may or may not be acceptable, depending on requirements. For example, if the void in the fillet did not extend through the entire braze width, the joint would still be leak tight, which was the major requirement of the brazement. On the other hand, if 100% braze fillet was needed because of stress requirements, the assembly would be unacceptable.

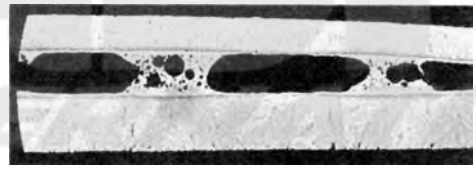


Fig. 17 Voids resulting from lack of fill between the faying surfaces of a lap joint between two sheets of Hastelloy X brazed with BNi-1 filler metal. Unetched. 16.5 ×. Source: Ref 1

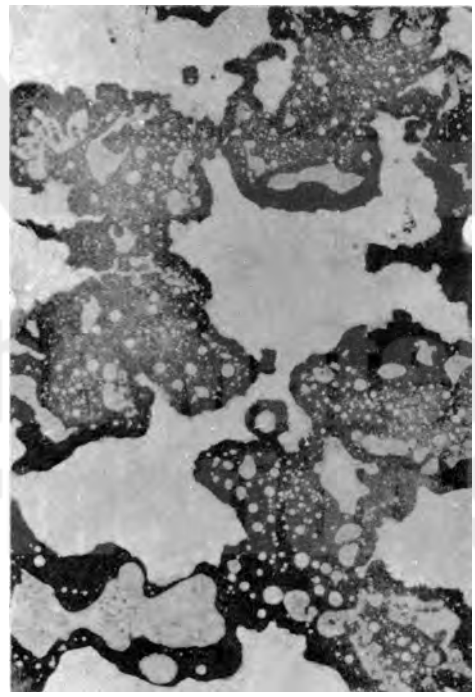


Fig. 18 Radiograph showing entrapped flux (dark areas) in a low carbon steel joint torch brazed with BAg-1 filler metal (light areas). 1×. Source: Ref 1

Base Metal Erosion. Certain brazing filler metals will readily alloy with the base metals being brazed, causing the constituents of the base metal to melt and, in some cases, creating an undercut condition or the actual disappearance of the faying surfaces. This is called base metal erosion. Extreme erosion in type 304 stainless steel brazed with a nickel-chromium-boron filler metal is shown in Fig. 20, and a similar joint without erosion is shown in Fig. 21. Erosion may not be serious where thick sections are to be joined, but it cannot be permitted where relatively thin sections are used.

Three factors influencing base metal erosion are brazing temperature, time at temperature, and the amount of brazing filler metal available or used in making the joint. As the brazing temperature exceeds the melting point of the filler metal, interaction between the molten filler metal and the base metal accelerates. Therefore, the brazing temperature should be kept low, provided, of course, that it is sufficient for proper flow of the filler

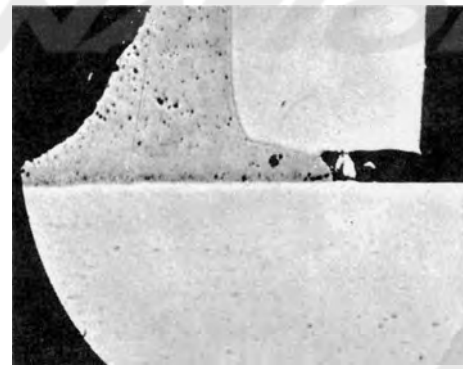


Fig. 19 Incomplete penetration of filler metal (BAg-1) in a brazed joint between copper components. 20×. Source: Ref 1

Fig. 20 Excessive erosion of type 304 stainless steel base metal by BNi-1 filler metal. Compare with the noneroded joint shown in Fig. 21. 20×. Source: Ref 1

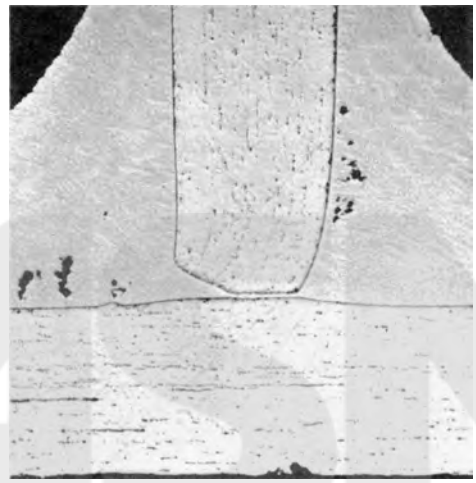


Fig. 21 Joint between type 304 stainless steel components brazed with BNi-1 filler metal, in which no base metal erosion occurred. Note characteristic sheared edge on one component and small voids in the filler metal. Source: Ref 1

metal to fill the joint. Similarly, time at temperature should be kept to a minimum to prevent excessive interaction between the molten filler metal and the base metal. Finally, the amount of filler metal required to fill the joint and provide the necessary fillet size should be closely controlled. Filler metal present in excess of the amount required is likely to react with the base metal, creating severe or excessive erosion in proportion to the amount of excess filler metal.

Joint Integrity

Some form of discontinuity usually occurs in all types of brazed joints. The degree and severity vary from a minor pinhole in the filler metal to gross discontinuities. Lack of fill or flux entrapment can vary from slight to nearly 100%. Erosion of the base metal can be nonexistent or can cause complete destruction of the joint.

Requirements for brazed joints are many and varied. As with other accepted joining processes, it is important that brazed joints be properly designed and engineered for the use intended. Significant factors involved are selection of proper base metals and brazing filler metal for compatibility and strength, proper fits and clearances, proper brazing process, and cleanliness of the surface to be brazed. Furthermore, it must be determined what requirements are necessary for withstanding the service conditions to which the finished brazement will be exposed. Primarily, brazed joints are designed for mechanical performance, electrical conductivity, or pressure tightness; therefore, the braze quality requirements should reflect the end use for which the joint was designed.

Methods of Inspection

Inspection of the completed assembly or subassembly is the last step in the brazing operation and is essential for ensuring satisfactory and uniform quality of the brazed unit. This inspection also provides a means for evaluating the adequacy of the design and the brazing operation with regard to ultimate integrity of the brazed unit.

Destructive methods such as peel tests, impact tests, torsion tests, and metallographic examination are initially used to determine whether the braze design meets the specified requirements. In production, these methods are employed only with random selection or lot testing of brazed joints. In lot testing, samples representing a small specified percentage of all production are tested to destruction. The results of these tests are assumed to apply to the entire production, and the joints in the various lots or batches are accepted or rejected accordingly.

When used as a check on an NDI method, such as visual examination, a production part can be selected at regular intervals and the joint tested to destruction so that rigid control of brazing procedures is maintained.

The inspection method chosen to evaluate the final brazed component should depend on service and reliability requirements. In many cases, the inspection methods are specified by the ultimate user or by regulatory codes. In establishing codes or specifications for brazed joints, the same approach should be used as in the setting up of standards for any other phase of manufacturing. The standards should be based, if possible, on requirements that have been established by prior service or history.

Visual Inspection

Visual inspection is the most widely used of the nondestructive methods for evaluating brazed joints. However, as with all other methods of inspection, visual inspection will not be effective if the joint cannot be readily examined. Visual inspection is also a convenient preliminary test where other inspection methods are used.

When brazing filler metal is fed from one side of the joint or replaced within the joint at or near one side so that visual examination of the opposite side of the joint after brazing shows a continuous fillet of filler metal, it can usually be assumed that the filler metal has flowed through the joint by capillary attraction and that a sound joint has been obtained.

On the other hand, if the joint can be inspected only on the side where the filler metal is applied, it is quite possible that an unsatisfactory joint has been produced, even though a satisfactory fillet is evident to the inspector.

Visual inspection cannot reveal internal discontinuities in a brazed joint that result from flux entrapment or lack of fill. Occasionally, gross erosion can be detected.

Proof Testing

Proof testing is a method of inspection that subjects the completed joints to loads slightly in excess of the loads to be applied during their subsequent service life. These loads can be applied by hydrostatic methods, tensile loading, spin testing, or numerous other methods. Occasionally, it is not possible to ensure a serviceable part by any of the other nondestructive methods of inspection, and proof testing then becomes the most satisfactory method.

Pressure Testing

Pressure testing of brazed assemblies is a method of leak testing and is usually confined to vessels and heat exchangers where liquid, gas, or air tightness is required. Several methods of pressure testing can be employed. Most use either air or gas, depending on the application of the vessel or heat exchanger. In most cases, the test pressures are greater than those to which the assembly will be subjected in service and are specified by the user.

One or more of the following three procedures are generally employed for pressure testing:

- All openings are closed. Air or gas is injected into the assembly until the specified pressure is reached. The inlet sources are closed off, the assembly is allowed to sit for a period of time, and pressure decreases are then measured on a gage
- All openings in the assembly are closed except one, which is fitted with an inlet pressure line. With the assembly submerged in a tank of water, air or gas is admitted through the inlet line until a specified pressure is reached. The inspector then looks for bubbles rising through the water
- All openings are closed, and the assembly is pressurized to the specified pressure. Then a leak detecting solution, of which there are several commercially available, is brushed on the joints to be inspected. If any of the joints leak, bubbling will occur

Vacuum and helium testing is generally used in inspecting assemblies where it is imperative that the most minute leak be detected. This method of inspection is often employed on nuclear reactor hardware. It is also extensively used in the inspection of refrigeration equipment. The assembly to be inspected is connected to a vacuum system, and the vacuum is monitored by a mass spectrometer.

Helium gas is flushed around the brazed joint; if any minute leak is present, the helium, because of its small molecule, will be pulled in by vacuum and register on the mass spectrometer, thus indicating the leak.

A more sensitive technique is pressurizing the assembly with helium while the assembly is contained in a sealed plastic bag. After pressurizing for a period of time (for example, 24 hours), the atmosphere in the bag is analyzed for the presence of helium.

Ultrasonic Inspection

Ultrasonic inspection, although not extensively used in the evaluation of brazed joints, can be the only method applicable in certain cases. The use of ultrasonic inspection depends largely on the design of the joint and the configuration of the adjacent areas of the brazed assembly. Advancements in ultrasonic inspection may increase the utility of this process so that brazed joints can be evaluated with reliability.

Radiographic Inspection

Radiographic inspection is commonly used for the nondestructive evaluation of brazed joints following visual examination. In almost all cases, the radiation beam is directed at about 90° to the plane of the joint, and the radiograph is taken through the thickness of the braze metal.

X-rays readily discern the differences in density between the brazing filler metal and the base metal. Care must be exercised because joints between sections of varying thicknesses can produce radiographs that are misleading and difficult to interpret. Also, it is often difficult to determine whether a joint has been penetrated fully or not at all; both situations yield radiographs in which there is a full fillet visible around the joint, and the gap in the joint itself has uniform radiographic density. By contrast, partly filled joints, voids in the braze metal, and inclusions are relatively easy to find with radiography.

The filler metal in brazed joints is very thin, from 0.013 to 0.25 mm (0.0005 to 0.010 in.) in thickness. When radiographs are made of brazed joints between thick components, the process may be unable to record the braze metal as a difference in density; at least 2% difference is usually needed for good sensitivity.

Liquid Penetrant Inspection

Liquid penetrant inspection is another nondestructive method for determining the reliability of brazed joints and assemblies. This inspection method produces a visual image of a discontinuity in the surface of the braze and reveals the nature of a discontinuity without impairing the parent metal. Acceptable and unacceptable components or assemblies can be separated in accordance with predetermined standards.

There are certain advantages obtained from the liquid penetrant inspection of brazed assemblies. However, a brazed joint or component should be visually inspected first, and then inspected by a liquid penetrant method

to resolve any doubt concerning joint integrity. Visual examination is restricted to those discontinuities that can be detected by the unaided eye. Liquid penetrant carries visual inspection a step further by increasing the detectability of fine cracks or openings.

Discontinuities such as incomplete fusion, cracks that x-rays cannot show because of orientation, and porosity and laps become visible with this technique. Liquid penetrants do not disclose subsurface discontinuities such as voids, cracks, or flux entrapment; radiography is best used to discover such discontinuities.

Selection of the specific liquid penetrant system for the inspection of brazed assemblies depends on the same factors as those that affect system selection for other workpieces. The water-washable, post-emulsifiable, and solvent-removable systems have been successfully used for inspecting brazed assemblies.

Inspection using liquid penetrants should not be performed prior to brazing unless adequate cleaning steps, such as vapor degreasing, are taken to remove entrapped penetrant fluid. If permitted to remain during the brazing cycle, this fluid can contaminate the furnace atmosphere and braze metal, producing flaws.

ACKNOWLEDGMENT

This chapter was adapted from Nondestructive Inspection of Weldments, Brazed Assemblies, and Soldered Joints, *Nondestructive Evaluation and Quality Control*, Volume 17, *ASM Handbook*, 1992.

REFERENCES

1. Nondestructive Inspection of Weldments, Brazed Assemblies, and Soldered Joints, *Nondestructive Evaluation and Quality Control*, Vol 17, *ASM Handbook*, ASM International, 1992, p 582–609

SELECTED REFERENCES

- R. Halmshaw, *Introduction to Non-Destructive Testing of Welded Joints*, 2nd ed., Woodhead Publishing Ltd., 1996

Index

A

B

C

D

E

F

G

H

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y

Z